

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:57:13 ; Search time 707.789 Seconds
(without alignments)
2110.838 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306362_306402
Perfect score: 41
Sequence: 1 ttgttgatgctgagcccc.....atactctattccgagtatga 41

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	611	5	BX489897
C 2	39.4	96.1	653	1	AI073549 ov45e08.x
3	37.8	92.2	705	5	BP160291
C 4	31.4	76.6	593	1	AA023625 mh80b06.f
5	31.4	76.6	616	4	BG084311 H3099A08-
6	31.4	76.6	627	2	BB200714 BB200714
7	31.4	76.6	2109	3	AK039911 Mus muscu
8	31.4	76.6	2591	3	AK077236 Mus muscu
9	31.4	76.6	2735	3	AK054182 Mus muscu
10	31.4	76.6	4248	3	AK041525 Mus muscu
11	31.4	76.6	4315	3	AK087638 Mus muscu
12	31.4	76.6	4321	3	AK036627 Mus muscu
13	29	70.7	365	1	AA092184 ll4609.se
C 14	28.8	70.2	576	1	AU041270 AU041270
15	27.8	67.8	656	9	CG483529 OST17199
16	23.4	57.1	1169	5	BU588298 AGENCOURT
C 17	23	56.1	871	8	BZ172878 CH230-505
18	22.6	55.1	748	9	AG356459 Mus muscu
19	22.4	54.6	213	2	AW582379 MR3-ST022
C 20	22.4	54.6	599	8	AZ833415 2M0115A06
C 21	22.4	54.6	646	9	CE331215 tigr-gss-
C 22	22	53.7	481	6	CA993426 PAR07E12
23	22	53.7	563	6	CA917333 EST641480
C 24	22	53.7	904	9	CG937676 MBEMW13TF

C 25	21.8	53.2	160	9	CG611862
26	21.8	53.2	399	5	BU874710
27	21.8	53.2	624	9	CE849311 tigr-gss-
28	21.8	53.2	628	7	CO023851 EST788984
29	21.8	53.2	637	6	CA116396 SCACLR105
30	21.8	53.2	673	6	CA100736 SCCCL700
31	21.8	53.2	725	6	CA261912 SCGLB201
32	21.6	52.7	220	4	BG145973 uu95b10.Y
33	21.6	52.7	275	2	BF922664 QV4-NT025
34	21.6	52.7	332	7	CO272045 EK105718.
C 35	21.6	52.7	492	5	BQ031691 UI-1-CF0-
C 36	21.6	52.7	510	6	CD291168 StrPu538.
C 37	21.6	52.7	687	7	CK134778 RH01453.3
C 38	21.6	52.7	968	5	BU957489 AGENCOURT
39	21.6	52.7	2219	3	AK047386 Mus muscu
40	21.4	52.2	467	8	AQ985326 RPCI-23-3
41	21.4	52.2	540	6	CD727728 4032587 1
C 42	21.4	52.2	543	7	CK981913 4114536 B
43	21.4	52.2	568	8	AZ858611 2M0163D19
C 44	21.4	52.2	589	4	BM090339 505541 MA
C 45	21.4	52.2	613	7	CK941792 4065309 B

ALIGNMENTS

RESULT 1
BX489897
LOCUS BX489897 611 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686N0562_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION BX489897
VERSION BX489897.1 GI:31998955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 611)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
TITLE ~~EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)~~
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp686N0562) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source

location/Qualifiers
1..611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N0562"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

ORIGIN

Query Match 100.0%; Score 41; DB 5; Length 611;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 41

```
Db      199 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 239
|||||
RESULT 2
AI073549/c
LOCUS
DEFINITION
ov45e08.xl Soares testis NHT Homo sapiens cDNA clone IMAGE:1640294
3', similar to gb:M12674 ESTROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION
AI073549
VERSION
AI073549.1 GI:3400193
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 653)
REFERENCE
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 483.
FEATURES
Location/Qualifiers
source
1..653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1640294"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis_NHT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 96.1%; Score 39.4; DB 1; Length 653;
Best Local Similarity 97.6%; Pred. No. 3.5e-05;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 458 TTGTTGGATGCTGAGCCCCCGATACTCTATTCCGAGTATGA 418
|||||
RESULT 3
BP160291
LOCUS
DEFINITION
BP160291 full-length enriched swine cDNA library, adult thymus Sus
scrofa cDNA clone THY010029A10 5', mRNA sequence.
ACCESSION
BP160291
VERSION
BP160291.1 GI:40409764
KEYWORDS
EST.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 593)
REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE
The WashU-HHMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 705)
REFERENCE
AUTHORS
Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
Okumura, N., Hamasima, N. and Awata, T.
TITLE
PEDE (Pig EST Data Explorer): construction of a database for ESTs
derived from porcine full-length cDNA libraries
Nucleic Acids Res. 32 (1), D484-D488 (2004)
CONTACT: Hirohide Uenishi
Animal Genome Laboratory, Genome Research Department
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
Tel: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in
Animal Genome Research Program (Japan) by National Institute of
Agrobiological Sciences and STAFF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA
library
Vector sequences were eliminated by RepeatMasker version 2002/07/13
and crossmatch version 0.990319
Low quality bases were trimmed based on the quality values.
FEATURES
Location/Qualifiers
source
1..705
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="THY010029A10"
/tissue_type="thymus"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult
thymus"
ORIGIN
Query Match 92.2%; Score 37.8; DB 5; Length 705;
Best Local Similarity 95.1%; Pred. No. 0.00016;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 553 TTGTTGGAGGCTGAGCCCCCAATCTATTCTCGAGTATGA 593
|||||
RESULT 4
AA023625/c
LOCUS
DEFINITION
mh80b06.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
clone IMAGE:457235 5' similar to gb:M12674 ESTROGEN RECEPTOR
(HUMAN); gb:M38651 Mouse estrogen receptor mRNA, complete cds
(MOUSE);, mRNA sequence.
ACCESSION
AA023625
VERSION
AA023625.1 GI:1487542
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 593)
REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE
The WashU-HHMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
```

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:274123

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 492.

Location/Qualifiers

FEATURES

source

1. .593
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:457235"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
/note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 76.6%; Score 31.4; DB 1; Length 593;
Best Local Similarity 85.4%; Pred. No. 0.063;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 41
|||||
Db 306 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 266
|||||

RESULT 5
BG084311
LOCUS
DEFINITION
H3099A08-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3099A08 5', mRNA sequence.
BG084311
BG084311.2 GI:40071640
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
10922068
On Jan 26, 2001 this sequence version replaced gi:12566875.
Other_ESTs: H3099A08-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3099 row: A column: 08
Seq primer: -21M13 Reverse
High quality sequence stop: 616
POLYA=No. Location/Qualifiers

FEATURES

source

1. .616
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3099A08-5"
/db_xref="taxon:10090"
/clone="H3099A08"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 76.6%; Score 31.4; DB 4; Length 616;
Best Local Similarity 85.4%; Pred. No. 0.063;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 41
|||||
Db 371 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 411
|||||

RESULT 6

BB200714

LOCUS

DEFINITION

BB200714.2

musculus cDNA clone A430027L01 3', mRNA sequence.

BB200714.2

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 627)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,

Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,

Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,

Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,

Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

On Jun 30, 2000 this sequence version replaced gi:8865667.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL:http://genome.gsc.riken.jp/

Unpublished (2001)

URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers		The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2591)	
FEATURES	source	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers	
	misc_feature	1. .2109 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:A430027L01" /db_xref="taxon:10090" /clone="A430027L01" /tissue_type="thymus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="0 day neonate" 1. .2109 /note="estrogen receptor 1 (alpha) (MGD MG1:1352467, GB NM_007956, evidence: BLASTN, 99%, match=1960)"	
ORIGIN	Query Match Best Local Similarity Matches	76.6%; 85.4%; 35;	Score 31.4; Pred. No. 0.071; Conservative 0;
			DB 3; Indels 6; Gaps 0;
	QY	1	TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
	Db	1652	TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 1692
RESULT 8	AK077236	2591 bp mRNA linear HTC 03-APR-2004	
LOCUS	Mus musculus 11 days pregnant adult female ovary and uterus cDNA, RIKEN full-length enriched library, clone:5031409B20		
DEFINITION	product:estrogen receptor 1 (alpha), full insert sequence.		
ACCESSION	AK077236		
VERSION	AK077236.1	GI:26097248	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
AUTHORS	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)		
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE	5 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
JOURNAL
MEDLINE
PUBMED
REFERENCE

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
JOURNAL
MEDLINE
PUBMED
REFERENCE

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL
MEDLINE
PUBMED
REFERENCE

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
11076861
JOURNAL
MEDLINE
PUBMED
REFERENCE

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
11042159
JOURNAL
MEDLINE
PUBMED
REFERENCE

6
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. .2735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

REFERENCE
source

/db_xref="FANTOM_DB:E230025M07"
/db_xref="taxon:10090"
/clone="E230025M07"
/sex="female"
/tissue type="oviduct"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="2 days pregnant adult"
213..1713
/note="estrogen receptor 1 (alpha) (MGD|MG1:1352467, GB|NM_007956, evidence: BLASTN, 99%, match=1960) putative"

misc_feature

ORIGIN

Query Match 76.6%; Score 31.4; DB 3; Length 2735;
Best Local Similarity 85.4%; Pred. No. 0.072;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCACTCTATTCCGAGTATGA 41
|||||
Db 1180 TTGTTGGATGCTGAACCGCCCATCATCTATTCTGAATATGA 1220
|||||

RESULT 10
AK041525 4248 bp mRNA linear HTC 03-APR-2004
LOCUS
DEFINITION
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630019F10 product:estrogen receptor 1 (alpha), full insert sequence.

ACCESSION
AK041525 GI:26334544
VERSION
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
JOURNAL
MEDLINE
PUBMED
REFERENCE

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
JOURNAL
MEDLINE
PUBMED
REFERENCE

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL
MEDLINE
PUBMED
REFERENCE

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
11076861
JOURNAL
MEDLINE
PUBMED
REFERENCE

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
11042159
JOURNAL
MEDLINE
PUBMED
REFERENCE

6
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. .2735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

REFERENCE
source

REFERENCE 6 (bases 1 to 4248)
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAKA, T., HIROZANE, T.,
HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KODA, M.,
KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSATO, N.,
OKAZAKI, Y., SAITO, R., SAITOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
SANO, H., SASAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T.,
SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKAHIRA, S.,
TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A.,
MURAMATSU, M. and HAYASHIZAKI, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
FEATURES
source
Location/Qualifiers
1..4248
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A630019F10"
/db_xref="taxon:10090"
/clone="A630019F10"
/tissue type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
331..2130
/note="unnamed protein product; estrogen receptor 1
(alpha) (MGD|MG1:1352467, GB|NM_007956, evidence: BLASTN,
99%, match=1960)
putative"
/codon_start=1
/protein_id="BAC30973.1"
/db_xref="GI:26334545"
/translation="MTMTLHTKASGMALLHQIQNELEPLNRPLQKMPMERALGEVYV
DNSKPTVFNYPEGAAYEFNAAAAAAASAPVYGSGIAYGPGSEAAAFSANSIGAFP
QLNSVSPSPMLLHPPLQLSPFLPHGQQVPYILENPSAYAVRDTGPPAFVRSNSDN
RRQNGRERLSSNEKGNIMESAKETRYCAVNCNVASGYHYGVWSCEGCKAFKRSIQ
GHNDYMCPATNQCITDKNRKSCQACRLKCYEVMKGGIRKDRRGGRLMKHRQRD
DLEGRNEMGASGDMRAANLWPSPLVIKHTKNSPALSLTADQMSALLDAEPPMIYSE
YDPSRPFSEASMGLLTNLADRELVMINWAKRVPFGFDLNLHDQVLLCEAWLEILM
IGLVWRSMEHPKLLFAPNLLDRNQCKVEGMVEIFDMLLATSSFRMMNLQGEFV
CLKSIILLNSGVYTFLSLKSEKDHIRVLDKITDTLIHLMAKAGLTLOQQRRL
AQLLILSHIRHMSNKGMEHLYNMCKNVVPLYDLLEMLDAHRLHAPASRMGVPEE
PSQTLATTSSTSAHSLQTYIYPPEAEQFPNTI"

Query Match 76.6%; Score 31.4; DB 3; Length 4248;
Best Local Similarity 85.4%; Pred. No. 0.075;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCGCCCACTCTATTCCGAGTATGA 41
|||||
Db 1297 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 1337
|||||

RESULT 11
AK087638
LOCUS
DEFINITION

AK087638 4315 bp mRNA linear HTC 03-APR-2004
Mus musculus 2 days pregnant adult female oviduct cDNA, RIKEN
full-length enriched library, clone:E230030D02 product:estrogen
receptor 1 (alpha), full insert sequence.

ACCESSION
AK087638
VERSION
GI:26104409
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4315)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT
cdna library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

FEATURES
source
Location/Qualifiers
1..4315
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:E230030D02"
/db_xref="taxon:10090"
/clone="E230030D02"
/sex="female"
/tissue_type="oviduct"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
1..4315
/note="estrogen receptor 1 (alpha) (MGD|MG1:1352467,
GB|NM_007956, evidence: BLASTN, 99%, match=1960)"

misc_feature
1..4315
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:E230030D02"
/db_xref="taxon:10090"
/clone="E230030D02"
/sex="female"
/tissue_type="oviduct"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"

ORIGIN
Query Match 76.6%; Score 31.4; DB 3; Length 4315;
Best Local Similarity 85.4%; Pred. No. 0.076;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 41
|||||
Db 1129 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGATATGA 1169
|||||

RESULT 12
AK036627
LOCUS
DEFINITION
Mus musculus adult male bone cDNA, RIKEN full-length enriched
library, clone:9830143G19 product:estrogen receptor 1 (alpha), full
insert sequence.
AK036627
ACCESSION
AK036627.1 GI:26331559
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913

11076861
4
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
JOURNAL
Functional annotation of a full-length mouse cDNA collection
REFERENCE
AUTHORS
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
JOURNAL
Analysis of the mouse transcriptome based on functional annotation
REFERENCE
AUTHORS
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4321)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cdna library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

FEATURES
source
Location/Qualifiers
1..4321
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:9830143G19"
/db_xref="taxon:10090"
/clone="9830143G19"
/sex="male"
/tissue_type="bone"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
183..1982
/note="unnamed protein product; estrogen receptor 1
(alpha) (MGD|MG1:1352467, GB|NM_007956, evidence: BLASTN,
99%, match=1960)
putative"
/codon_start=1
/protein_id="BAC29510.1"
/db_xref="GI:26331560"
/translation="MTWTLHTKASGMALLHQIQNELEPLNRPQLKMPMERALGEVYV
DNSKPTVFNYPEGAAAYEFNAAAAAASAPVYQSGIAYGPGSEAAAFSANSLGAFP
QLNSVSPPLMLLHPPQLSPFLPHGQQVPYYLENEPSAYAVRDTGPPAFYRSNSDN
RRQNGRERLSSSNEKGNIMESAKETRYCAVCNDYASGYHYGVMSCEGCKAFFKRSIQ
GHNDYCPATNQCTIDKNRRKSCQACRLRKCVEGMKGGIRKDRRGGRMLKHKRQRD
DLEGRNEMGASGDMRAANLWPSPLVIKTKNSPALSLTADQMVSALLDAEPPMIYSE
YDPSRPFSEASMGLLTNLADRELVHMINKRVPGFGDLNHDQVHLLLECWLLEILM
IGLVWRSMEHPGKLLFAPNLLLDNRQCKVEGMVEIFDMLLATSSRFRMNLQGEFV
CLKSIILLNSGVYFSLSTLKSLEEXDHIHRVLDKITDTLLHLMKAGLTIQQQHRRL
AQLLLILSHIRHMSNKGMEHLYNMCKNVVPLDYDLLLEMLDAHRLHAPASRMGVPEE
PSQTLATTSSTSAHSLQTYIIPPEAGFPNTI"
4306..4311
polyA_signal


```
polyA_site      /note="putative"
4321
/note="putative"

ORIGIN
Query Match      76.6%; Score 31.4; DB 3; Length 4321;
Best Local Similarity 85.4%; Pred. No. 0.076;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 1149 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 1189

RESULT 13
AA092184
LOCUS
DEFINITION
114609.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
cDNA 5', mRNA sequence.
ACCESSION
AA092184
VERSION
AA092184.1 GI:1637173
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 365)
AUTHORS
Liew,C.C.
TITLE
cDNAs from fetal heart (1996)
JOURNAL
Unpublished (1996)
COMMENT
Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6179750995
Email: cliew@rics.bwh.harvard.edu
PCR Primers
FORWARD: 5' GCCAAGCTCGAATTAAACCCCTCACTAAAGGG 3'
BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGCG 3'
Seq primer: 5' GAAATTAACCCCTCACTAAAGGG 3'.
Location/Qualifiers
1..365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/lab_host="E. coli XL1-Blue"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/note="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."

ORIGIN
Query Match      70.7%; Score 29; DB 1; Length 365;
Best Local Similarity 97.6%; Pred. No. 0.57;
Matches 40; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 114 TTGTTGGATGCTGAG-CCCCCATACTCTATTCCGAGTATGA 153

RESULT 14
AU041270/c
LOCUS
DEFINITION
AU041270 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
J0828F12 3', mRNA sequence.
ACCESSION
AU041270
VERSION
AU041270.1 GI:3955505
KEYWORDS
EST.

polyA_site      /note="putative"
4321
/note="putative"

ORIGIN
Query Match      76.6%; Score 31.4; DB 3; Length 4321;
Best Local Similarity 85.4%; Pred. No. 0.076;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 1149 TTGTTGGATGCTGAACCGCCCATGATCTATTCTGAATATGA 1189

RESULT 13
AA092184
LOCUS
DEFINITION
114609.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
cDNA 5', mRNA sequence.
ACCESSION
AA092184
VERSION
AA092184.1 GI:1637173
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 365)
AUTHORS
Liew,C.C.
TITLE
cDNAs from fetal heart (1996)
JOURNAL
Unpublished (1996)
COMMENT
Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6179750995
Email: cliew@rics.bwh.harvard.edu
PCR Primers
FORWARD: 5' GCCAAGCTCGAATTAAACCCCTCACTAAAGGG 3'
BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGCG 3'
Seq primer: 5' GAAATTAACCCCTCACTAAAGGG 3'.
Location/Qualifiers
1..365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/lab_host="E. coli XL1-Blue"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/note="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."

ORIGIN
Query Match      70.7%; Score 29; DB 1; Length 365;
Best Local Similarity 97.6%; Pred. No. 0.57;
Matches 40; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 114 TTGTTGGATGCTGAG-CCCCCATACTCTATTCCGAGTATGA 153

RESULT 14
AU041270/c
LOCUS
DEFINITION
AU041270 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
J0828F12 3', mRNA sequence.
ACCESSION
AU041270
VERSION
AU041270.1 GI:3955505
KEYWORDS
EST.
```

```
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
REFERENCE
1 (bases 1 to 576)
AUTHORS
Ko,M.S.H., Kitchen,J.R., Wang,X., Threat,T.A., Sun,T.,
Depalma,G.E., Liang,Y., Kargul,G.J., Sharara,R., Lim,M.K. and
Doi,H.
TITLE
Systematic analyses of genes expressed in 4-cell mouse embryo (The
ERATO/Doi Project at Wayne State University)
JOURNAL
Unpublished (1998)
COMMENT
Contact: Hirofumi Doi
Doi Bioasymmetry Project, ERATO
Japan Science and Technology Corporation (JST)
WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
Email: hd@bioa.jst.go.jp.
Location/Qualifiers
1..576
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="J0828F12"
/dev_stage="four-cell-embryo"
/clone_lib="Mouse four-cell-embryo cDNA"

ORIGIN
Query Match      70.2%; Score 28.8; DB 1; Length 576;
Best Local Similarity 80.5%; Pred. No. 0.72;
Matches 33; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 41
|||||
Db 460 TTGTNGAATGCTGAACCGCCCATGATCTATTCTGAATATGA 420

RESULT 15
CG483529
LOCUS
DEFINITION
CG483529 Mus musculus 129Sv/Ev Mus musculus genomic clone OST17199,
genomic survey sequence.
ACCESSION
CG483529
VERSION
CG483529.1 GI:37238317
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
REFERENCE
1 (bases 1 to 656)
AUTHORS
Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
TITLE
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT
Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
Location/Qualifiers
1..656
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"

SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
REFERENCE
1 (bases 1 to 656)
AUTHORS
Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
TITLE
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT
Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
Location/Qualifiers
1..656
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
```

/clone="OST17199"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 67.8%; Score 27.8; DB 9; Length 656;
Best Local Similarity 80.0%; Pred. No. 1.9;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATG 40
|||||
Db 457 TTGTTGNATGCTGAACCGCCCATGATCTATTCTGAATATG 496
|||||

Search completed: November 3, 2004, 09:43:33
Job time : 712.123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 06:51:04 ; Search time 74.6435 Seconds
(without alignments)
2816.475 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306362_306402

Perfect score: 41

Sequence: 1 ttgttgatgctgagcccc.....atactctattccgagtatga 41

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	392	14	US-10-052-092-4
2	41	100.0	392	15	US-10-437-107-4
3	41	100.0	1218	15	US-10-027-983-98
4	41	100.0	1218	15	US-10-448-753-98
5	41	100.0	1237	14	US-10-052-092-3
6	41	100.0	1237	15	US-10-437-107-3
7	41	100.0	1245	15	US-10-157-899A-1
8	41	100.0	1245	15	US-10-157-899A-5
9	41	100.0	1245	15	US-10-157-899A-7
10	41	100.0	1245	15	US-10-157-899A-9
11	41	100.0	1317	15	US-10-157-899A-3
12	41	100.0	1317	15	US-10-157-899A-11

13	41	100.0	1317	15	US-10-157-899A-13
14	41	100.0	1317	15	US-10-157-899A-15
15	41	100.0	1542	15	US-10-006-760-18
16	41	100.0	1644	14	US-10-052-092-2
17	41	100.0	1644	15	US-10-437-107-2
18	41	100.0	1785	15	US-10-095-373A-1
19	41	100.0	1788	9	US-09-853-033-1
20	41	100.0	1983	9	US-09-853-033-3
21	41	100.0	1983	9	US-09-853-033-5
22	41	100.0	1983	9	US-09-853-033-7
23	41	100.0	2028	15	US-10-095-373A-72
24	41	100.0	2031	15	US-10-095-373A-68
25	41	100.0	2092	10	US-09-952-680A-9
26	41	100.0	2092	14	US-10-052-092-6
27	41	100.0	2092	15	US-10-157-899A-54
28	41	100.0	2092	15	US-10-437-107-6
29	41	100.0	2106	14	US-10-052-092-5
30	41	100.0	2106	15	US-10-437-107-5
31	41	100.0	2178	15	US-10-095-373A-70
32	41	100.0	2181	15	US-10-095-373A-66
33	41	100.0	4963	9	US-09-281-674-9
34	41	100.0	4963	9	US-09-777-317-9
35	41	100.0	4963	9	US-09-892-227-9
36	41	100.0	6450	13	US-10-096-710-2
37	41	100.0	6450	13	US-10-081-563-1
38	41	100.0	6450	14	US-10-052-092-1
39	41	100.0	6450	14	US-10-052-092-7
40	41	100.0	6450	15	US-10-207-655-60
41	41	100.0	6450	15	US-10-177-293-127
42	41	100.0	6450	15	US-10-027-983-3
43	41	100.0	6450	15	US-10-007-926A-422
44	41	100.0	6450	15	US-10-437-107-1
45	41	100.0	6450	15	US-10-437-107-7

ALIGNMENTS

RESULT 1

US-10-052-092-4
; Sequence 4, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-4

Query Match 100.0%; Score 41; DB 14; Length 392;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATCTCTATTCGAGTATGA 41

Db 182 TTGTTGGATGCTGAGCCCCCATCTCTATTCGAGTATGA 222

RESULT 2

```
US-10-437-107-4
; Sequence 4, Application US/10437107
; Publication No. US20030186313A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/437,107
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US/10/052,092
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Human
US-10-437-107-4

Query Match      100.0%; Score 41; DB 15; Length 392;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      182 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 222

RESULT 3
US-10-027-983-98
; Sequence 98, Application US/10027983
; Publication No. US20030139360A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 98
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-027-983-98

Query Match      100.0%; Score 41; DB 15; Length 1218;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425

RESULT 4
US-10-448-753-98
; Sequence 98, Application US/10448753
; Publication No. US20030211611A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/448,753
```

```
US-10-437-107-3
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US/10/027,983
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 98
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-448-753-98

Query Match      100.0%; Score 41; DB 15; Length 1218;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425

RESULT 5
US-10-052-092-3
; Sequence 3, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-3

Query Match      100.0%; Score 41; DB 14; Length 1237;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      544 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 584

RESULT 6
US-10-437-107-3
; Sequence 3, Application US/10437107
; Publication No. US20030186313A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/437,107
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US/10/052,092
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
```



```
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Human
US-10-437-107-3
```

```
Query Match          100.0%; Score 41; DB 15; Length 1237;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
        |||||||
Db      544 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 584
```

RESULT 7

```
US-10-157-899A-1
; Sequence 1, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of the human-Estrogen-receptor-ligand-binding-domain and a
; OTHER INFORMATION: zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1245)
US-10-157-899A-1
```

```
Query Match          100.0%; Score 41; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
        |||||||
Db      406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446
```

RESULT 8

```
US-10-157-899A-5
; Sequence 5, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
```

```
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; OTHER INFORMATION: and a zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1245)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (613)..(615)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (721)..(723)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (733)..(735)
US-10-157-899A-5
```

```
Query Match          100.0%; Score 41; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
        |||||||
Db      406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446
```

RESULT 9

```
US-10-157-899A-7
; Sequence 7, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; OTHER INFORMATION: and a zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1245)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (712)..(714)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (733)..(735)
US-10-157-899A-7
```

```
Query Match          100.0%; Score 41; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
      |||||||
Db      406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446

RESULT 10
US-10-157-899A-9
; Sequence 9, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; and a zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1245)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1021)..(1023)
US-10-157-899A-9

Query Match          100.0%; Score 41; DB 15; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
      |||||||
Db      406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446

RESULT 11
US-10-157-899A-3
; Sequence 3, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 3
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of the human-Estrogen-receptor-ligand-binding-domain and a
; OTHER INFORMATION: zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1317)
US-10-157-899A-3

Query Match          100.0%; Score 41; DB 15; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
      |||||||
Db      478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 12
US-10-157-899A-11
; Sequence 11, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; and a zinc finger array (C7)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1317)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (685)..(687)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (793)..(795)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (805)..(807)
US-10-157-899A-11

Query Match          100.0%; Score 41; DB 15; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
      |||||||
Db      478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 13
US-10-157-899A-13
```

```
; Sequence 13, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; OTHER INFORMATION: and a zinc finger array(C7)
; NAME/KEY: CDS
; LOCATION: (1)..(1317)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (784)..(786)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (805)..(808)
; US-10-157-899A-13

Query Match      100.0%; Score 41; DB 15; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 14
US-10-157-899A-15
; Sequence 15, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion of a mutated human-Estrogen-receptor-ligand-binding-domain
; OTHER INFORMATION: and a zinc finger array(C7)
; NAME/KEY: CDS
; LOCATION: (1)..(1317)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (784)..(786)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (805)..(808)
; US-10-157-899A-13

Query Match      100.0%; Score 41; DB 15; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 15
US-10-006-760-18
; Sequence 18, Application US/10006760
; Publication No. US20030186385A1
; GENERAL INFORMATION:
; APPLICANT: Koide, Shohei
; TITLE OF INVENTION: METHOD OF IDENTIFYING POLYPEPTIDE MONOBODIES WHICH BIND
; TITLE OF INVENTION: TO TARGET PROTEINS AND USE THEREOF
; FILE REFERENCE: 176/60901
; CURRENT APPLICATION NUMBER: US/10/006,760
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/249,756
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 18
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: lexA-ER(alpha)EF fusion protein
; US-10-006-760-18

Query Match      100.0%; Score 41; DB 15; Length 1542;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      703 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 743

Search completed: November 3, 2004, 10:01:03
Job time : 74.6435 secs
```

```
; NAME/KEY: CDS
; LOCATION: (1)..(1317)
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1093)..(1095)
; US-10-157-899A-15

Query Match      100.0%; Score 41; DB 15; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 15
US-10-006-760-18
; Sequence 18, Application US/10006760
; Publication No. US20030186385A1
; GENERAL INFORMATION:
; APPLICANT: Koide, Shohei
; TITLE OF INVENTION: METHOD OF IDENTIFYING POLYPEPTIDE MONOBODIES WHICH BIND
; TITLE OF INVENTION: TO TARGET PROTEINS AND USE THEREOF
; FILE REFERENCE: 176/60901
; CURRENT APPLICATION NUMBER: US/10/006,760
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/249,756
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 18
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: lexA-ER(alpha)EF fusion protein
; US-10-006-760-18

Query Match      100.0%; Score 41; DB 15; Length 1542;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      703 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 743

Search completed: November 3, 2004, 10:01:03
Job time : 74.6435 secs
```

This Page Blank (uspto)
BEST AVAILABLE COPY

Season 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:44:23 ; Search time 83.8636 Seconds
(without alignments)
2566.384 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306362_306402

Perfect score: 41

Sequence: 1 ttgttgatgctgagcccc.....atactctatccgagtatga 41

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	100.0	392	6	ABK89698	Abk89698 Oestrogen
2	41	100.0	1044	6	ABZ23390	Abz23390 Nucleotid
3	41	100.0	1044	6	ABZ23391	Abz23391 Reverse c
4	41	100.0	1218	10	ADB81478	ADB81478 Human ESR
5	41	100.0	1223	4	AAF29924	Aaf29924 Human est
6	41	100.0	1237	6	ABK89697	Abk89697 Oestrogen
7	41	100.0	1245	8	AAD53878	Aad53878 C7LBDAS c
8	41	100.0	1245	8	AAD53879	Aad53879 C7LBDAS c
9	41	100.0	1245	8	AAD53875	Aad53875 C7LBDAS c
10	41	100.0	1245	8	AAD53877	Aad53877 C7LBDAS c
11	41	100.0	1317	8	AAD53876	Aad53876 C7LBDAS c
12	41	100.0	1317	8	AAD53882	Aad53882 C7LBDAS c
13	41	100.0	1317	8	AAD53881	Aad53881 C7LBDAS c
14	41	100.0	1317	8	AAD53880	Aad53880 C7LBDAS c
15	41	100.0	1374	4	AAF29925	Aaf29925 Human est
16	41	100.0	1380	4	AAC86919	Aac86919 Nucleotid
17	41	100.0	1446	5	AAC88199	Aac88199 Mammalian
18	41	100.0	1542	10	ADE39221	Ade39221 LexA-oest
19	41	100.0	1644	6	ABK89696	Abk89696 Oestrogen
20	41	100.0	1770	10	ADB99351	Adb99351 Fusion pr
21	41	100.0	1785	10	AAD64536	Aad64536 Human wil

22	41	100.0	1788	6	ABL57497	AbL57497 Human nuc
23	41	100.0	1803	10	ADB99349	Adb99349 Fusion pr
24	41	100.0	1956	2	AAV17756	Aav17756 Multi-chi
25	41	100.0	1983	6	ABL57500	AbL57500 Cre recom
26	41	100.0	1983	6	ABL57499	AbL57499 Cre recom
27	41	100.0	1983	6	ABL57498	AbL57498 Cre recom
28	41	100.0	2028	10	AAD64590	Aad64590 Human ER3
29	41	100.0	2031	10	AAD64588	Aad64588 Human ER3
30	41	100.0	2092	1	AAN70880	Aan70880 cDNA enco
31	41	100.0	2092	4	AAC86920	Aac86920 Nucleotid
32	41	100.0	2092	6	ABK89700	Abk89700 Oestrogen
33	41	100.0	2092	6	ABL51900	AbL51900 Human oes
34	41	100.0	2092	8	AAD53909	Aad53909 Human wil
35	41	100.0	2092	10	ADC09969	Adc09969 Human oes
36	41	100.0	2106	6	ABK89699	Abk89699 Oestrogen
37	41	100.0	2178	10	AAD64589	Aad64589 Human ER3
38	41	100.0	2181	10	AAD64587	Aad64587 Human ER3
39	41	100.0	2220	6	ABS56328	Abs56328 Human oes
40	41	100.0	2220	6	ABA01104	AbA01104 Fused ER
41	41	100.0	2220	10	ADC51161	Adc51161 Human oes
42	41	100.0	2220	10	ADC51169	Adc51169 Variant h
43	41	100.0	2220	10	ADC51171	Adc51171 Variant h
44	41	100.0	2220	10	ADC51170	Adc51170 Variant h
45	41	100.0	2322	2	AAx60628	Aax60628 SSR-LBD f

ALIGNMENTS

RESULT 1
ABK89698
ID ABK89698 standard; DNA; 392 BP.
XX
AC ABK89698;
XX

05-NOV-2002 (first entry)

Oestrogen receptor alpha nucleic acid comprising A908G mutation, #4.

Oestrogen receptor alpha; breast cancer; pre-malignant lesion;
invasive breast cancer; A908G oestrogen receptor alpha transition;
oestrogen receptor alpha K303R substitution; human; gene; ds.

Homo sapiens.

WO200257283-A1.

25-JUL-2002.

16-JAN-2002; 2002WO-US004982.

19-JAN-2001; 2001US-0262990P.

09-JUL-2001; 2001US-0304018P.

(BAYU) BAYLOR COLLEGE MEDICINE.

Fuqua S, O'connell P, Allred DC, Hopp TA;

WPI; 2002-590711/63.

New isolated estrogen receptor alpha with A908G mutation or K303R substitution, useful as diagnostic marker in breast tissue such as pre-malignant lesions for the development of breast cancer, particularly invasive breast cancer.

Claim 1; Page; 133pp; English.

The invention relates to an isolated oestrogen receptor alpha nucleic acid sequence comprising an A908G mutation, or an amino acid sequence comprising a K303R substitution. Also described are methods for detecting susceptibility to development of breast cancer or invasive breast cancer in an individual, for diagnosing breast cancer in an individual; and for screening for a modulator of an oestrogen receptor alpha polypeptide

comprising a K303R substitution. The oestrogen receptor alpha is useful as a diagnostic marker in breast tissue such as pre- malignant lesions for the development of breast cancer, particularly invasive breast cancer. The methods are useful for determining susceptibility to development of breast cancer, for diagnosing, preventing or treating breast cancer. Transgenic mice may be used for screening and identifying agents that interact with the oestrogen receptor alpha, or affect breast tissue health. The A908G oestrogen receptor alpha transition is frequently present in pre-malignant lesions of the breast and can occur in the adjacent normal-appearing breast epithelium. The present sequence represents an oestrogen receptor alpha nucleic acid which may comprise the A908G mutation. Note: The present sequence is not shown in the specification but was obtained by the indexer from GenBank using the accession number given in the specification

Sequence 392 BP; 97 A; 92 C; 124 G; 79 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 6; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 41
182 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 222

RESULT 2
ABZ23390

ID ABZ23390 standard; cDNA; 1044 BP.

AC ABZ23390;

07-APR-2003 (first entry)

Nucleotide sequence of human oestrogen receptor alpha bait protein.

Human; oestrogen receptor alpha; ER-alpha; CF16; CF17; CF18; CF19; CF40; CF41; CF42; CF43; cofactor; osteoporosis; bone disease; reproduction; hormonal dysfunction; cancer; cardiovascular disease; atherosclerosis; hot flush; mood change; Alzheimer's disease; gene; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1044
/*tag= a
/product= "oestrogen receptor alpha bait"

WO200270699-A2.

12-SEP-2002.

28-FEB-2002; 2002WO-EP002189.

01-MAR-2001; 2001EP-00105062.

(LION-) LION BIOSCIENCE AG.

Albers M, Ellwanger S, Loeser E, Koegl M;
WPI; 2002-713451/77.

New cofactors of estrogen receptor alpha, designated as CF16, CF17, CF18, CF19, CF40, CF41, CF42 and/or CF43, useful for screening of compounds for treating osteoporosis, hormonal dysfunctions, cancer or cardiovascular diseases.

Disclosure; Page 108-109; 111pp; English.

The present sequence encodes a human oestrogen receptor alpha (ER-alpha) bait protein. The specification describes cofactors of ER-alpha, designated CF16, CF17, CF18, CF19, CF40, CF41, CF42, and CF43. The

cofactor polypeptides and nucleic acid molecules are useful for screening for compounds for treating osteoporosis and other bone diseases, failures in reproductive functions or hormonal dysfunctions, cancer or cardiovascular diseases such as atherosclerosis, and in preventing hot flushes, mood changes and Alzheimer's disease. The CF proteins are also useful for screening for ligands of the ER alpha. The nucleic acid sequences are useful for making vectors and CF polypeptides, transforming host cells, as research tools for developing nucleic acid probes, and for developing analytical tools such as antisense oligonucleotides

Sequence 1044 BP; 252 A; 272 C; 295 G; 225 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 6; Length 1044;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 41
211 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 251

RESULT 3
ABZ23391/c

ID ABZ23391 standard; cDNA; 1044 BP.

AC ABZ23391;

07-APR-2003 (first entry)

Reverse complement of human oestrogen receptor alpha bait protein cDNA.

Human; oestrogen receptor alpha; ER-alpha; CF16; CF17; CF18; CF19; CF40; CF41; CF42; CF43; cofactor; osteoporosis; bone disease; reproduction; hormonal dysfunction; cancer; cardiovascular disease; atherosclerosis; hot flush; mood change; Alzheimer's disease; ss.

Homo sapiens.

WO200270699-A2.

12-SEP-2002.

28-FEB-2002; 2002WO-EP002189.

01-MAR-2001; 2001EP-00105062.

(LION-) LION BIOSCIENCE AG.

Albers M, Ellwanger S, Loeser E, Koegl M;
WPI; 2002-713451/77.

New cofactors of estrogen receptor alpha, designated as CF16, CF17, CF18, CF19, CF40, CF41, CF42 and/or CF43, useful for screening of compounds for treating osteoporosis, hormonal dysfunctions, cancer or cardiovascular diseases.

Disclosure; Page 109-110; 111pp; English.

The present sequence represents the reverse complement of an oestrogen receptor alpha (ER-alpha) bait protein. The specification describes cofactors of ER-alpha, designated CF16, CF17, CF18, CF19, CF40, CF41, CF42, and CF43. The cofactor polypeptides and nucleic acid molecules are useful for screening for compounds for treating osteoporosis and other bone diseases, failures in reproductive functions or hormonal dysfunctions, cancer or cardiovascular diseases such as atherosclerosis, and in preventing hot flushes, mood changes and Alzheimer's disease. The CF proteins are also useful for screening for ligands of the ER alpha. The nucleic acid sequences are useful for making vectors and CF polypeptides, transforming host cells, as research tools for developing nucleic acid probes, and for developing analytical tools such as antisense oligonucleotides

SQ Sequence 1044 BP; 225 A; 295 C; 272 G; 252 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 6; Length 1044;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 834 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 794
|||||

RESULT 4
ADB81478
ID ADB81478 standard; DNA; 1218 BP.
XX
AC ADB81478;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human ESR-alpha with fatty acid synthase (ESR-alpha FASN) DNA.
XX
KW human; ds; oestrogen receptor alpha; ESR-alpha FASN; fatty acid synthase;
KW oestrogen receptor 1; ESR1; NR3A1; bone maintenance;
KW cardiovascular system; cancer; gene therapy; hyperproliferative disease;
KW inflammation; tumour formation; infection; cytostatic; antiinflammatory;
KW antimicrobial.
XX
OS Homo sapiens.
XX
PN WO2003052072-A2.
XX
PD 26-JUN-2003.
XX
PF 13-DEC-2002; 2002WO-US040083.
XX
PR 18-DEC-2001; 2001US-00027983.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Dobie KW, Roach MP;
XX
DR WPI; 2003-577322/54.
XX
PT New antisense compound targeted to nucleic acid encoding estrogen
PT receptor alpha and inhibiting expression of estrogen receptor alpha,
PT useful for treating a disease or condition e.g. a hyperproliferative
PT disease.
XX
PS Example 17; Page 232; 232pp; English.
XX
CC This invention relates to human oestrogen receptor alpha (ESR-alpha), and
CC the novel antisense oligonucleotides that modulate its expression. The
CC oestrogen receptor alpha protein is also known as oestrogen receptor 1,
CC ESR1, and NR3A1. Oestrogen, the steroid hormone ligand of ESR-alpha, is
CC important for bone maintenance and plays a protective role in the
CC cardiovascular system, as well as being required for normal sexual
CC maturation through promoting growth and differentiation. Splice variants
CC of ESR-alpha, however, have been associated with various cancers
CC including the breast and pituitary. Accordingly, antisense
CC oligonucleotides that inhibit the expression of ESR-alpha in cells or
CC tissues can be used in gene therapy to treat conditions such as
CC hyperproliferative disease, inflammation, tumour formation and to prevent
CC or delay infection. As such, the present invention describes these
CC antisense oligos as having cytostatic, antiinflammatory and antimicrobial
CC activities. This polynucleotide is the DNA of the human oestrogen
CC receptor alpha splice variant with fatty acid synthase designated ESR-
CC alpha FASN of the invention.
XX
SQ Sequence 1218 BP; 261 A; 367 C; 328 G; 262 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 10; Length 1218;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425
|||||

RESULT 5
AAF29924
ID AAF29924 standard; DNA; 1223 BP.
XX
AC AAF29924;
XX
DT 04-APR-2001 (first entry)
XX
DE Human estrogen receptor alpha isoform #1 DNA.
XX
KW Human; estrogen receptor alpha; cancer; osteoporosis; bone; Alzheimer's;
KW cardiovascular; ds.
XX
OS Homo sapiens.
XX
PN WO200100823-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-EP005981.
XX
PR 29-JUN-1999; 99IT-MI001433.
XX
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX
PI Gannon F, Dengger S, Flouriot G;
XX
DR WPI; 2001-137955/14.
XX
PT Novel isoforms of human estrogen receptor alpha useful for preparing
PT therapeutic agents for treating cancer, osteoporosis, Alzheimer's disease
PT and cardiovascular diseases.
XX
PS Claim 1; Pag e44; 53pp; English.
XX
CC The present invention relates to a human estrogen receptor (hER)-alpha
CC isoform. Molecules which modulate the activity of the estrogen receptor
CC are useful for the preparation of therapeutic agents for treating cancer,
CC osteoporosis and other bone disorders, Alzheimer's disease and
CC cardiovascular diseases
XX
SQ Sequence 1223 BP; 303 A; 316 C; 341 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 4; Length 1223;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 390 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 430
|||||

RESULT 6
ABK89697
ID ABK89697 standard; DNA; 1237 BP.
XX
AC ABK89697;
XX
DT 05-NOV-2002 (first entry)
XX
DE Oestrogen receptor alpha nucleic acid comprising A908G mutation, #3.
XX
KW Oestrogen receptor alpha; breast cancer; pre-malignant lesion;
KW invasive breast cancer; A908G oestrogen receptor alpha transition;
KW oestrogen receptor alpha K303R substitution; human; gene; ds.
XX
OS Homo sapiens.

XX WO200257283-A1.
XX 25-JUL-2002.
XX 16-JAN-2002; 2002WO-US004982.
XX 19-JAN-2001; 2001US-0262990P.
XX 09-JUL-2001; 2001US-0304018P.
XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX Fuqua S, O'connell P, Allred DC, Hopp TA;
XX WPI; 2002-590711/63.
XX New isolated estrogen receptor alpha with A908G mutation or K303R
XX substitution, useful as diagnostic marker in breast tissue such as pre-
XX malignant lesions for the development of breast cancer, particularly
XX invasive breast cancer.
XX Claim 1; Page; 133pp; English.
XX The invention relates to an isolated oestrogen receptor alpha nucleic
XX acid sequence comprising an A908G mutation, or an amino acid sequence
XX comprising a K303R substitution. Also described are methods for detecting
XX susceptibility to development of breast cancer or invasive breast cancer
XX in an individual, for diagnosing breast cancer in an individual; and for
XX screening for a modulator of an oestrogen receptor alpha polypeptide
XX comprising a K303R substitution. The oestrogen receptor alpha is useful
XX as a diagnostic marker in breast tissue such as pre- malignant lesions
XX for the development of breast cancer, particularly invasive breast
XX cancer. The methods are useful for determining susceptibility to
XX development of breast cancer, for diagnosing, preventing or treating
XX breast cancer. Transgenic mice may be used for screening and identifying
XX agents that interact with the oestrogen receptor alpha, or affect breast
XX tissue health. The A908G oestrogen receptor alpha transition is
XX frequently present in pre-malignant lesions of the breast and can occur
XX in the adjacent normal-appearing breast epithelium. The present sequence
XX represents an oestrogen receptor alpha nucleic acid which may comprise
XX the A908G mutation. Note: The present sequence is not shown in the
XX specification but was obtained by the indexer from GenBank using the
XX accession number given in the specification
XX Sequence 1237 BP; 369 A; 218 C; 286 G; 364 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 41; DB 6; Length 1237;
XX Best Local Similarity 100.0%; Pred. No. 2.5e-07;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 544 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 584
RESULT 7
AAD53878
ID AAD53878 standard; DNA; 1245 BP.
XX AAD53878;
XX 28-MAY-2003 (first entry)
XX C7LBDAS chimeric mutant DNA #2.
XX Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.
XX Homo sapiens.
OS Unidentified.
OS Chimeric.

XX Key Location/Qualifiers
FH 1. .1245
FT /*tag= a
FT /product= "C7LBDAS fusion protein"
FT replace(712,A)
FT /*tag= b
FT mutation replace(733. .735, CTG)
FT /*tag= c
XX WO200297050-A2.
XX 05-DEC-2002.
XX 31-MAY-2002; 2002WO-US016946.
XX 31-MAY-2001; 2001US-0294839P.
XX (NOVS) NOVARTIS AG.
XX Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
XX Zerby DB;
XX WPI; 2003-156794/15.
XX P-PSDB; AAE35278.
XX New mutant estrogen receptor ligand binding domain capable of interacting
XX with non-endogenous ligand, useful e.g. in combination with a ligand for
XX constructing selective molecular gene switches for regulating gene
XX function.
XX Example 4; Page 126-127; 159pp; English.
XX The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
XX binding domain (LBD) which comprises an amino acid modification in region
XX 1, region 2 or both and interacting with a non-endogenous ligand as a
XX result of the amino acid modification. Sequences of the invention are
XX useful for treatment of genetic diseases, acquired diseases and any other
XX conditions including cell proliferative disorders such as cancer e.g.
XX lung, breast, lymphoid, gastrointestinal, genito-urinary tract
XX adenocarcinomas and other malignancies such as colon cancers, renal- cell
XX carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
XX of the small intestine and cancer of the oesophagus. The invention is
XX useful for constructing selective molecular gene switches for regulating
XX gene function in plants and transgenic animals. It is also useful in gene
XX therapy. The present sequence is human ER alpha LBD -zinc finger array
XX (C7) chimeric mutant DNA
XX Sequence 1245 BP; 281 A; 365 C; 328 G; 271 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 41; DB 8; Length 1245;
XX Best Local Similarity 100.0%; Pred. No. 2.5e-07;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446
RESULT 8
AAD53879
ID AAD53879 standard; DNA; 1245 BP.
XX AAD53879;
XX 28-MAY-2003 (first entry)
XX C7LBDAS chimeric mutant DNA #3.
XX Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.

XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1245
FT /*tag= a
FT /product= "C7LBDAS fusion protein"
FT mutation replace(1021..1023, CAT)
FT /*tag= b

WO200297050-A2.

XX
XX
PD 05-DEC-2002.

XX 31-MAY-2002; 2002WO-US016946.

XX 31-MAY-2001; 2001US-0294839P.

XX (NOVS) NOVARTIS AG.

XX Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;

XX WPI; 2003-156794/15.
DR P-PSDB; AAE35279.

XX
PT New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.

XX
PS Example 4; Page 129-131; 159pp; English.

XX
CC The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric mutant DNA

XX SQ Sequence 1245 BP; 281 A; 364 C; 329 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446

RESULT 9
AAD53875
ID AAD53875 standard; DNA; 1245 BP.

XX
AC AAD53875;

XX 28-MAY-2003 (first entry)

XX C7LBDAS chimeric DNA #1.

XX Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;

KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; gene; ds.

XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.

XX Key Location/Qualifiers
FT CDS 1..1245
FT /*tag= a
FT /product= "C7LBDAS fusion protein"

XX WO200297050-A2.

XX
XX
PD 05-DEC-2002.

XX 31-MAY-2002; 2002WO-US016946.

XX 31-MAY-2001; 2001US-0294839P.

XX (NOVS) NOVARTIS AG.

XX Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;

XX WPI; 2003-156794/15.
DR P-PSDB; AAE35275.

XX
PT New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.

XX
PS Example 4; Page 115-117; 159pp; English.

XX
CC The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric DNA

XX SQ Sequence 1245 BP; 282 A; 364 C; 327 G; 272 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 406 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 446

RESULT 10
AAD53877
ID AAD53877 standard; DNA; 1245 BP.

XX
AC AAD53877;

XX 28-MAY-2003 (first entry)

XX C7LBDAS chimeric mutant DNA #1.

KW Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;

KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1245
FT /*tag= a
FT /product= "C7LBDA5 fusion protein"
FT replace(613,A)
FT /*tag= b
FT mutation
FT replace(721..722, AT)
FT /*tag= c
FT mutation
FT replace(733..735, CTG)
FT /*tag= d
XX
PN WO200297050-A2.
XX
XX
PD 05-DEC-2002.
XX
XX
PF 31-MAY-2002; 2002WO-US016946.
XX
XX
PR 31-MAY-2001; 2001US-0294839P.
XX
XX (NOVS) NOVARTIS AG.
XX
XX Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;
PI
XX
DR WPI; 2003-156794/15.
DR P-PSDB; AAE35277.
XX
XX New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.
XX
XX
PS Example 4; Page 122-124; 159pp; English.
XX
XX The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric mutant DNA
XX
SQ Sequence 1245 BP; 281 A; 365 C; 328 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 41
|||||
Db 406 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 446

RESULT 11
AAD53876
ID AAD53876 standard; DNA; 1317 BP.
XX
XX AAD53876;
AC
XX

DT 28-MAY-2003 (first entry)
XX
DE C7LBDBS chimeric DNA #1.
XX
KW Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; gene; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1317
FT /*tag= a
FT /product= "C7LBDBS fusion protein"
XX
PN WO200297050-A2.
XX
XX
PD 05-DEC-2002.
XX
XX
PF 31-MAY-2002; 2002WO-US016946.
XX
XX
PR 31-MAY-2001; 2001US-0294839P.
XX
XX (NOVS) NOVARTIS AG.
XX
XX Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;
PI
XX
DR WPI; 2003-156794/15.
DR P-PSDB; AAE35276.
XX
XX New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.
XX
XX
PS Example 4; Page 118-120; 159pp; English.
XX
XX The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric DNA
XX
SQ Sequence 1317 BP; 305 A; 372 C; 361 G; 279 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 41
|||||
Db 478 TTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGA 518

RESULT 12
AAD53882
ID AAD53882 standard; DNA; 1317 BP.
XX
XX AAD53882;
AC
XX

DT 28-MAY-2003 (first entry)
XX
DE C7LBDBS chimeric mutant DNA #3.
XX
KW Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1317
FT /*tag= a
FT /product= "C7LBDBS fusion protein"
FT replace(1093..1095, CAT)
FT /*tag= b
XX
PN WO200297050-A2.
XX
PD 05-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US016946.
XX
PR 31-MAY-2001; 2001US-0294839P.
XX
PA (NOVS) NOVARTIS AG.
XX
PI Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;
XX
DR WPI; 2003-156794/15.
DR P-PSDB; AAE35282.
XX
PT New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.
XX
PS Example 4; Page 140-142; 159pp; English.
XX
CC The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric mutant DNA
XX
SQ Sequence 1317 BP; 304 A; 372 C; 363 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

RESULT 13
AAD53881
ID AAD53881 standard; DNA; 1317 BP.
XX

AC AAD53881;
XX
DT 28-MAY-2003 (first entry)
XX
DE C7LBDBS chimeric mutant DNA #2.
XX
KW Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1317
FT /*tag= a
FT /product= "C7LBDBS fusion protein"
FT replace(784,A)
FT /*tag= b
FT mutation
FT mutation
FT mutation
FT mutation
FT mutation
FT mutation
XX
PN WO200297050-A2.
XX
PD 05-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US016946.
XX
PR 31-MAY-2001; 2001US-0294839P.
XX
PA (NOVS) NOVARTIS AG.
XX
PI Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;
XX
DR WPI; 2003-156794/15.
DR P-PSDB; AAE35281.
XX
PT New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.
XX
PS Example 4; Page 136-138; 159pp; English.
XX
CC The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric mutant DNA
XX
SQ Sequence 1317 BP; 304 A; 373 C; 362 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 8; Length 1317;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518

```
RESULT 14
AAD53880
ID AAD53880 standard; DNA; 1317 BP.
XX AC
AC AAD53880;
XX
DT 28-MAY-2003 (first entry)
XX
DE C7LBDBS chimeric mutant DNA #1.
XX
KW Oestrogen receptor; ER alpha; ligand binding domain; genetic disease;
KW acquired disease; cell proliferative disorder; cancer; adenocarcinoma;
KW LBD; gene switch; transgenic animal; transgenic; gene therapy; human;
KW chimeric; zinc finger array; C7; mutant; gene; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1317
FT /*tag= a
FT /product= "C7LBDBS fusion protein"
FT mutation replace(685,A)
FT /*tag= b
FT mutation replace(793..794, AT)
FT /*tag= c
FT mutation replace(805..807, CTG)
FT /*tag= d
XX
PN WO200297050-A2.
XX
XX 05-DEC-2002.
XX
XX 31-MAY-2002; 2002WO-US016946.
XX
PR 31-MAY-2001; 2001US-0294839P.
XX
PA (NOVS ) NOVARTIS AG.
XX
PI Bracken KR, De Los Angeles JE, Huang Y, Kadan MJ, Ksander GM;
PI Zerby DB;
XX
DR WPI; 2003-156794/15.
DR P-PSDB; AAE35280.
XX
PT New mutant estrogen receptor ligand binding domain capable of interacting
PT with non-endogenous ligand, useful e.g. in combination with a ligand for
PT constructing selective molecular gene switches for regulating gene
PT function.
XX
PS Example 4; Page 133-134; 159pp; English.
XX
CC The invention relates to a mutant oestrogen receptor (ER) alpha-ligand
CC binding domain (LBD) which comprises an amino acid modification in region
CC 1, region 2 or both and interacting with a non-endogenous ligand as a
CC result of the amino acid modification. Sequences of the invention are
CC useful for treatment of genetic diseases, acquired diseases and any other
CC conditions including cell proliferative disorders such as cancer e.g.
CC lung, breast, lymphoid, gastrointestinal, genito-urinary tract
CC adenocarcinomas and other malignancies such as colon cancers, renal- cell
CC carcinoma, prostate cancer, non-small cell carcinoma of the lung, cancer
CC of the small intestine and cancer of the oesophagus. The invention is
CC useful for constructing selective molecular gene switches for regulating
CC gene function in plants and transgenic animals. It is also useful in gene
CC therapy. The present sequence is human ER alpha LBD -zinc finger array
CC (C7) chimeric mutant DNA
XX
SQ Sequence 1317 BP; 304 A; 373 C; 362 G; 278 T; 0 U; 0 Other;
```

Query Match 100.0%; Score 41; DB 8; Length 1317;
Best Local Similarity 100.0%; Pred. NO. 2.6e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 478 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 518
|||||

RESULT 15
AAF29925
ID AAF29925 standard; DNA; 1374 BP.
XX
AC AAF29925;
XX
DT 04-APR-2001 (first entry)
XX
DE Human estrogen receptor alpha isoform #2 DNA.

XX
KW Human; estrogen receptor alpha; cancer; osteoporosis; bone; Alzheimer's;
KW cardiovascular; ds.

XX OS Homo sapiens.
XX WO200100823-A1.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-EP005981.
XX PR 29-JUN-1999; 99IT-MI001433.
XX PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX Gannon F, Denger S, Flouriot G;
XX WPI; 2001-137955/14.

XX PT Novel isoforms of human estrogen receptor alpha useful for preparing
PT therapeutic agents for treating cancer, osteoporosis, Alzheimer's disease
PT and cardiovascular diseases.

XX PS Claim 1; Page 44-45; 53pp; English.

XX CC The present invention relates to a human estrogen receptor (hER)-alpha
CC isoform. Molecules which modulate the activity of the estrogen receptor
CC are useful for the preparation of therapeutic agents for treating cancer,
CC osteoporosis and other bone disorders, Alzheimer's disease and
CC cardiovascular diseases

XX SQ Sequence 1374 BP; 349 A; 346 C; 386 G; 293 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. NO. 2.6e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 541 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 581
|||||

Search completed: November 3, 2004, 06:59:25
Job time : 84.8636 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 05:58:29 ; Search time 18.8325 Seconds
(without alignments)
1547.447 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306362_306402
Perfect score: 41
Sequence: 1 ttgttgatgctgagcccc.....atactattccgagtatga 41

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	1218	4	US-10-027-983-98 Sequence 98, Appl
2	41	100.0	1956	3	US-08-693-940-2 Sequence 2, Appli
3	41	100.0	1956	4	US-09-566-660-2 Sequence 2, Appli
4	41	100.0	2322	3	US-08-564-264-2 Sequence 2, Appli
5	41	100.0	4963	1	US-08-076-726-16 Sequence 16, Appl
6	41	100.0	4963	1	US-08-260-452-9 Sequence 9, Appli
7	41	100.0	4963	2	US-08-481-970-9 Sequence 9, Appli
8	41	100.0	4963	2	US-08-897-719-9 Sequence 9, Appli
9	41	100.0	4963	3	US-09-163-269-9 Sequence 9, Appli
10	41	100.0	4963	4	US-09-281-674-9 Sequence 9, Appli
11	41	100.0	6450	3	US-09-041-886-34 Sequence 34, Appl
12	41	100.0	6450	4	US-08-453-998-1 Sequence 1, Appli
13	41	100.0	6450	4	US-10-027-983-3 Sequence 3, Appli
14	41	100.0	8566	4	US-10-027-983-10 Sequence 10, Appl
15	21.6	52.7	334	4	US-09-270-767-8774 Sequence 8774, Ap
16	21.6	52.7	334	4	US-09-270-767-24056 Sequence 24056, A
17	21	51.2	976	3	US-08-961-083-79 Sequence 79, Appl
18	21	51.2	976	4	US-09-536-784-79 Sequence 79, Appl
19	21	51.2	1929	4	US-09-583-110-1148 Sequence 1148, Ap
20	21	51.2	11384	4	US-08-961-527-45 Sequence 45, Appl
21	20.4	49.8	3182	4	US-09-221-017B-1029 Sequence 1029, Ap
22	20	48.8	1230025	4	US-09-198-452A-1 Sequence 1, Appli
23	19.8	48.3	9458	4	US-08-956-171E-114 Sequence 114, App
24	19.8	48.3	9458	4	US-08-781-986A-114 Sequence 114, App
25	19.8	48.3	17041	1	US-08-076-011-1 Sequence 1, Appli
26	19.4	47.3	432	4	US-09-513-999C-2913 Sequence 2913, Ap
27	19.4	47.3	202001	4	US-09-734-674-3 Sequence 3, Appli

C 28	19.2	46.8	1134	4	US-09-270-767-11368	Sequence 11368, A
C 29	19.2	46.8	35081	2	US-08-752-760A-1	Sequence 1, Appli
30	19	46.3	1000	4	US-09-641-638-460	Sequence 460, App
31	19	46.3	1000	4	US-10-170-097-460	Sequence 460, App
C 32	19	46.3	1809	4	US-09-489-039A-36	Sequence 36, Appl
33	19	46.3	4266	4	US-09-489-039A-7	Sequence 7, Appli
34	18.8	45.9	252	4	US-09-313-294A-2329	Sequence 2329, Ap
C 35	18.8	45.9	446	4	US-09-513-999C-1069	Sequence 1069, Ap
36	18.8	45.9	1104	4	US-09-328-352-3352	Sequence 3352, Ap
C 37	18.8	45.9	319608	4	US-09-539-333D-1	Sequence 1, Appli
C 38	18.8	45.9	319608	4	US-09-679-409-1	Sequence 1, Appli
39	18.6	45.4	271	4	US-09-513-999C-8712	Sequence 8712, Ap
40	18.6	45.4	1242	3	US-09-068-132-1	Sequence 1, Appli
41	18.6	45.4	1242	3	US-09-068-132-2	Sequence 2, Appli
42	18.6	45.4	1242	4	US-10-018-730A-1	Sequence 1, Appli
43	18.6	45.4	3501	4	US-09-645-593-4	Sequence 4, Appli
44	18.6	45.4	9897	4	US-08-961-527-10	Sequence 10, Appl
C 45	18.6	45.4	17425	3	US-09-511-625B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-027-983-98
; Sequence 98, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 98
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-027-983-98

Query Match 100.0%; Score 41; DB 4; Length 1218;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 41
Db 385 TTGTTGGATGCTGAGCCCCCATCTCTATTCCGAGTATGA 425

RESULT 2
US-08-693-940-2
; Sequence 2, Application US/08693940
; Patent No. 6133027
; GENERAL INFORMATION:
; APPLICANT: Yee, Jiing-Kuan
; APPLICANT: Friedman, Theodore
; APPLICANT: Chen, Shin-Tai
; TITLE OF INVENTION: Inducible Expression System
; TITLE OF INVENTION: Useful in the Generation of Packaging Cell Lines for
; TITLE OF INVENTION: Pseudotyped Retroviral Vectors
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible


```
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,940
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Francis, Carol L
; REGISTRATION NUMBER: 36,513
; REFERENCE/DOCKET NUMBER: 6510-055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1956 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1953
; OTHER INFORMATION:
;
US-08-693-940-2

Query Match      100.0%; Score 41; DB 3; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      1123 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1163

RESULT 3
US-09-566-660-2
; Sequence 2, Application US/09566660
; Patent No. 6432705
; GENERAL INFORMATION:
; APPLICANT: Yee, Jiing-Kuan
; APPLICANT: Friedmann, Theodore
; APPLICANT: Chen, Shin-Tai
; TITLE OF INVENTION: Inducible Expression System
; FILE REFERENCE: 6510-055CON
; CURRENT APPLICATION NUMBER: US/09/566,660
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 08/693,940
; PRIOR FILING DATE: 1996-08-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1953)
; OTHER INFORMATION: Multi-chimeric transactivating factor
;
US-09-566-660-2

Query Match      100.0%; Score 41; DB 4; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      1123 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1163
```

```
RESULT 4
US-08-564-264-2
; Sequence 2, Application US/08564264
; Patent No. 6040430
; GENERAL INFORMATION:
; APPLICANT: STEWART, Francis
; TITLE OF INVENTION: REGULATION OF SITE-SPECIFIC RECOMBINATION BY SITE-SPECIFIC RECOMBINASE/NUCLEAR
; TITLE OF INVENTION: RECOMBINATION BY SITE-SPECIFIC RECOMBINASE/NUCLEAR
; TITLE OF INVENTION: RECEPTOR FUSION PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,264
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02088
; FILING DATE: 28-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 110 298.2
; FILING DATE: 28-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-5019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1-1269)
; OTHER INFORMATION: /note= "FLP recombinase domain."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1270-1284)
; OTHER INFORMATION: /note= "Linker peptide."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1285-2322)
; OTHER INFORMATION: /note= "Estrogen binding domain."
;
US-08-564-264-2

Query Match      100.0%; Score 41; DB 3; Length 2322;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      1489 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1529

RESULT 5
US-08-076-726-16
; Sequence 16, Application US/08076726
; Patent No. 5464758
; GENERAL INFORMATION:
```

```

; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Tight Control of Gene Expression in
; TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,726
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-076-726-16

Query Match 100.0%; Score 41; DB 1; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1460

RESULT 6
US-08-260-452-9
; Sequence 9, Application US/08260452
; Patent No. 5650298
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
; TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,452
; FILING DATE:

; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
; TITLE OF INVENTION: Transcription
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,970
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/260,452
; FILING DATE: 14-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; FILING DATE: 14-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: pUHD BGR4
; US-08-260-452-9

Query Match 100.0%; Score 41; DB 1; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1460

RESULT 7
US-08-481-970-9
; Sequence 9, Application US/08481970
; Patent No. 5859310
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
; TITLE OF INVENTION: Transcription
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,970
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/260,452
; FILING DATE: 14-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; FILING DATE: 14-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Human cytomegalovirus
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE: PUHD BGR4
CLONE: PUHD BGR4
US-08-481-970-9

Query Match 100.0%; Score 41; DB 2; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||
Db 1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1460

RESULT 8
US-08-897-719-9
; Sequence 9, Application US/08897719
; Patent No. 5922927
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
; TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,719
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/260,452
; FILING DATE:
; APPLICATION NUMBER: 08/076,327
; FILING DATE: 14-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: PUHD BGR4
; US-08-897-719-9

Query Match 100.0%; Score 41; DB 2; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||
Db 1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1460

RESULT 9
US-09-163-269-9
; Sequence 9, Application US/09163269
; Patent No. 6252136
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,269
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/481,970
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; FILING DATE: 14-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: PUHD BGR4
; US-09-163-269-9

Query Match 100.0%; Score 41; DB 3; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||
Db 1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1460

RESULT 10

```
US-09-281-674-9
; Sequence 9, Application US/09281674
; Patent No. 6783756
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
;           Bujard, Hermann
;           Salfeld, Jochen
;           Voss, Jeffrey
; TITLE OF INVENTION: Methods for Regulating Gene Expression
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,674
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,306
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/260,452
; FILING DATE: 14-JUN-1994
; APPLICATION NUMBER: 08/076,327
; FILING DATE: 14-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: pUHD BGR4
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-09-281-674-9
Query Match      100.0%; Score 41; DB 4; Length 4963;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCGAGTATGA 41
Db      1420 TTGTTGGATGCTGAGCCCCCATACTCTATTCGAGTATGA 1460

RESULT 11
US-09-041-886-34
; Sequence 34, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 361..2146
; US-09-041-886-34

Query Match      100.0%; Score 41; DB 3; Length 6450;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCGAGTATGA 41
Db      1315 TTGTTGGATGCTGAGCCCCCATACTCTATTCGAGTATGA 1355

RESULT 12
US-08-453-998-1
; Sequence 1, Application US/08453998
; Patent No. 6444438
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, PIERRE
; APPLICANT: METZGER, DANIEL
; APPLICANT: WHITE, JOHN
; TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
; TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND
; TITLE OF INVENTION: CORRESPONDING TRANSFORMED STRAINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVE., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,998
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```



```
; APPLICATION NUMBER: US 07/161,064
; FILING DATE: 03-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CHAPIN, MARLANA K.
; REGISTRATION NUMBER: 35,843
; REFERENCE/DOCKET NUMBER: 1037/98493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-453-998-1

Query Match      100.0%; Score 41; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      1315 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1355

RESULT 13
US-10-027-983-3
; Sequence 3, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (361)...(2148)
US-10-027-983-3

Query Match      100.0%; Score 41; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      1315 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1355

RESULT 14
US-10-027-983-10
; Sequence 10, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 10
; LENGTH: 8566
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(2302)
; NAME/KEY: exon:exon junction
; LOCATION: (1491)...(1492)
; OTHER INFORMATION: exon 4:exon 5
; NAME/KEY: start codon
; LOCATION: (2302)...(2305)
; NAME/KEY: 3'UTR
; LOCATION: (2735)...(8566)
; NAME/KEY: exon:exon junction
; LOCATION: (3569)...(3570)
; OTHER INFORMATION: exon 8:exon 9
; NAME/KEY: exon:exon junction
; LOCATION: (3708)...(3709)
; OTHER INFORMATION: exon 9:exon 10
; NAME/KEY: exon:exon junction
; LOCATION: (3842)...(3843)
; OTHER INFORMATION: exon 10:exon 11
; NAME/KEY: exon:exon junction
; LOCATION: (4026)...(4027)
; OTHER INFORMATION: exon 11:exon 12
US-10-027-983-10

Query Match      100.0%; Score 41; DB 4; Length 8566;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db      3431 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 3471

RESULT 15
US-09-270-767-8774
; Sequence 8774, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8774
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8774

Query Match      52.7%; Score 21.6; DB 4; Length 334;
Best Local Similarity 75.0%; Pred. No. 2.8;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      2 TGTGATGCTGAGCCCCCATACTCTATTCCGAGT 37
Db      31 TGTGAGCGGAGCCCACTACCCAAAGTTCGAGT 66

Search completed: November 3, 2004, 09:46:52
Job time : 20.8325 secs
```

Sequence

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:57:13 ; Search time 707.789 Seconds
(without alignments)
2110.838 Million cell updates/sec

Title: US-09-933-267A-1_COPY_64311_64351
Perfect score: 41
Sequence: 1 tctctcctgactttgttcgt.....aattgataagatctaatttg 41

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25.2	61.5	651	CA832730	MCS009G07
2	24.8	60.5	711	AI725341	1207 PtIF
3	24	58.5	641	AV400698	AV400698
4	23.4	57.1	540	BU779065	SJEEUA05
5	23.4	57.1	643	AZ381999	1M0138J17
6	23	56.1	449	BH234952	MHAA_D12.
C 7	23	56.1	798	CD765542	EST00228
C 8	22.8	55.6	499	AL921300	AL921300
9	22.8	55.6	633	CC063019	fgma002d0
C 10	22.8	55.6	711	BP678983	BP678983
11	22.8	55.6	1123	CL106907	ISB1-4711
C 12	22.8	55.6	1144	CL108627	ISB1-50C2
13	22.6	55.1	600	AZ771688	1M0574H06
C 14	22.6	55.1	630	BH181335	018_M_12-
C 15	22.6	55.1	630	CNS07MZF	T7_end of
16	22.6	55.1	657	CE675920	tigr-gss-
17	22.6	55.1	835	CC321221	TAM32-20L
18	22.6	55.1	1226	CG770712	TCB51.1_B
C 19	22.4	54.6	153	AA120330	mp85g05.r
C 20	22.4	54.6	179	AI790018	ue66f07.r
C 21	22.4	54.6	201	W12072	ma96c01.r1
C 22	22.4	54.6	208	AW143047	EST293342
C 23	22.4	54.6	221	AA467358	ve21a04.r
C 24	22.4	54.6	241	W12311	ma57d07.r1

C 25	22.4	54.6	250	1	AA929314	VZ40g07.r
C 26	22.4	54.6	250	1	AA177935	mt02d02.r
C 27	22.4	54.6	279	8	BH791401	SALK_0598
C 28	22.4	54.6	280	5	BX638156	BX638156
C 29	22.4	54.6	281	4	BM493236	EST00019
C 30	22.4	54.6	286	4	BM403207	EST00001
C 31	22.4	54.6	287	1	AI575934	UI-R-Y0-u
C 32	22.4	54.6	288	1	AA198597	mul5b02.r
33	22.4	54.6	291	1	AV084080	AV084080
34	22.4	54.6	294	2	BE134239	ug08e05.x
C 35	22.4	54.6	296	2	BE631235	uu58a01.y
C 36	22.4	54.6	300	7	CR470354	CR470354
C 37	22.4	54.6	302	7	W11844	mb21d07.r1
C 38	22.4	54.6	306	7	W35903	ma81c11.r1
C 39	22.4	54.6	311	7	W62151	md86h06.r1
C 40	22.4	54.6	312	2	AW988519	ug08e05.y
C 41	22.4	54.6	314	7	W54541	md09h10.r1
C 42	22.4	54.6	319	1	AA060917	mj80c09.r
C 43	22.4	54.6	319	1	AA241978	mw23c12.r
C 44	22.4	54.6	319	7	W62065	md84e01.r1
45	22.4	54.6	320	2	BF524186	UI-R-Y0-u

ALIGNMENTS

RESULT 1
CA832730/c
LOCUS CA832730 651 bp mRNA linear EST 12-DEC-2002
DEFINITION MCS009G07 150738 Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2 AM). Mesembryanthemum crystallinum cDNA clone MCS009G07 5, mRNA sequence.
ACCESSION CA832730
VERSION CA832730.1 GI:26560495
KEYWORDS EST.
SOURCE Mesembryanthemum crystallinum (common iceplant)
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 651)
AUTHORS Cushman, J.C.
TITLE An expressed sequence tag database for the common ice plant, Mesembryanthemum crystallinum
JOURNAL Unpublished (1997)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 009 row: G column: 07
Seq primer: T3 20mer
High quality sequence stop: 651.
Location/Qualifiers
1. .651
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="MCS009G07"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2 AM)."
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library construction was performed according to Stratagene's recommended protocol for the

ORIGIN		Lambda UniZapXR vector and cDNA synthesis kit."										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV400698										
VERSION		AV400698.1 GI:6904350										
KEYWORDS		EST.										
SOURCE		Bombyx mori (domestic silkworm)										
ORGANISM		Bombyx mori										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV400698										
VERSION		AV400698.1 GI:6904350										
KEYWORDS		EST.										
SOURCE		Bombyx mori (domestic silkworm)										
ORGANISM		Bombyx mori										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV400698										
VERSION		AV400698.1 GI:6904350										
KEYWORDS		EST.										
SOURCE		Bombyx mori (domestic silkworm)										
ORGANISM		Bombyx mori										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV400698										
VERSION		AV400698.1 GI:6904350										
KEYWORDS		EST.										
SOURCE		Bombyx mori (domestic silkworm)										
ORGANISM		Bombyx mori										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV400698										
VERSION		AV400698.1 GI:6904350										
KEYWORDS		EST.										
SOURCE		Bombyx mori (domestic silkworm)										
ORGANISM		Bombyx mori										
Query Match		61.5%; Score 25.2; DB 6; Length 651;										
Best Local Similarity		78.9%; Pred. No. 39;										
Matches		30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;										
QY	Db	4 CTCCTGACTTTGTCGTAATAATTGATAAGATCTAATTTG 41										
QY	Db	214 CTCAAGAGTTTGCAGATAAATTGACAAGGTCTAATTTG 177										
RESULT 2		AI725341 711 bp mRNA linear EST 11-JUN-1999										
LOCUS		AI725341 1207 PtIFG2 Pinus taeda cDNA clone 9300r, mRNA sequence.										
DEFINITION		AI725341										
ACCESSION		AI725341.1 GI:5044160										
VERSION		EST.										
KEYWORDS		Pinus taeda (loblolly pine)										
SOURCE		Pinus taeda										
ORGANISM		Pinus taeda										
REFERENCE	AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
		Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.										
TITLE	JOURNAL	1 (bases 1 to 711)										
		Kinlaw,C.S.										
COMMENT		Loblolly pine cDNAs										
		Unpublished (1995)										
USDA IFG Dendrome Project		Contact: Claire S. Kinlaw										
Institute of Forest Genetics		USDA IFG Dendrome Project										
Dendrome Project, Institute of Forest Genetics		Berkely, CA 94701										
Tel: 5105596429		Fax: 5105596440										
Email: csk@s27w007.pswfs.gov		Seq primer: M13 Reverse.										
Location/Qualifiers		1. 711										
source		/organism="Pinus taeda"										
		/mol_type="mRNA"										
		/db_xref="taxon:3352"										
		/clone="9300r"										
		/tissue_type="xylem"										
		/clone_lib="PtIFG2"										
		/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;										
		The tissue source for this library is xylem. The xylem										
		tissue was harvested in spring and summer from branches										
		of seed orchard trees which are clones of the same										
		genotype. Branches were 4-6 inches in diameter. The cDNAs										
		were directionally cloned into Lambda Zap and were										
		rescued as a Bluescript derivative in the EcoRI and XhoI										
		sites."										
ORIGIN		Query Match 60.5%; Score 24.8; DB 1; Length 711;										
Best Local Similarity		74.4%; Pred. No. 57;										
Matches		29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	Db	1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAATT 39										
QY	Db	449 TNTCCNCTGAATTTGTTNCTAAATTGAAAGGATCTTATT 487										
RESULT 3		AV400698 641 bp mRNA linear EST 05-FEB-2000										
LOCUS		AV400698 Bombyx mori brain Daizo P0 (just after pupation) Bombyx										
DEFINITION		mori cDNA clone br--1665 T3, mRNA sequence.										
ACCESSION		AV4										

TITLE Sequence characterization of an embryonic cDNA library from the
turkey (Meleagris gallopavo)
JOURNAL Unpublished (2003)
COMMENT Contact: Reed KM
Department of Veterinary Pathobiology.
University of Minnesota
295 AS/VM, 1988 Fitch Ave, St. Paul, MN 55108, USA
Tel: 612 624 1287
Fax: 612 625 0204
Email: reedx054@tc.umn.edu
Seq primer: TCGAAGTTCCCTCACTAAAGG
POLYA=Yes.

FEATURES
source
1. .798
/organism="Meleagris gallopavo"
/mol_type="mRNA"
/db_xref="taxon:9103"
/clone="7KRA07"
/sex="unknown"
/tissue_type="whole embryo"
/dev_stage="24-day"
/clone_lib="Turkey Lambda ZAP Library"
/note="Vector: Lambda ZAP Library ZAP; RNA was isolated
from a 24-day turkey embryo using a commercially available
kit (Stratagene. cDNA library was constructed using a
Lambda ZAP (SK-, Stratagene). "

ORIGIN
Query Match 56.1%; Score 23; DB 6; Length 798;
Best Local Similarity 74.4%; Pred. No. 2.9e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 TCTCCTGACTTTGTCCGTAAATTGATAGATCTAATTG 41
||||| ||| ||||||| ||||||| |||||

Db 348 TCTCTTGTAATTTCGGTAAATGCATTAGATCAACTTG 310

RESULT 8
AL921300/c
LOCUS AL921300 PUR-Z1+Z2 499 bp mRNA linear EST 06-JUL-2004
DEFINITION AL921300 Danio rerio cDNA clone 093-G06-2, mRNA sequence.
ACCESSION AL921300
VERSION AL921300.1 GI:23186598
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 499)
Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., He, Y., Ma, W.,
Wang, W., Wen, Z. and Peng, J.
15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis
Genome Res. 13 (3), 455-466 (2003)
22505427
12618376
Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengjr@imcb.a-star.edu.sg
Clone requests: info@openbiosystems.com
Open Biosystems,
6705 Odyssey Drive, Huntsville, AL 35806.

FEATURES
source
1. .499
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="093-G06-2"

/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PUR-Z1+Z2"

ORIGIN
Query Match 55.6%; Score 22.8; DB 1; Length 499;
Best Local Similarity 79.4%; Pred. No. 3.3e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTCTCCTGACTTTGTCCGTAAATTCATAAGATCT 35
||||| ||||||| ||||||| |||||

Db 146 CTCTCCAACTTTGTTCAGTAAAGTTGAGTAAATCT 113

RESULT 9
CC063019 633 bp DNA linear GSS 15-APR-2003
LOCUS fgma002d003f11t0 fgma filtered library (LibID: 145) Glycine max
DEFINITION genomic clone fgma002d003f11, genomic survey sequence.
ACCESSION CC063019
VERSION CC063019.1 GI:29842401
KEYWORDS GSS.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 633)
Budiman, M.A., Freese, R.G., Flick, E., Citek, R.W., Nunberg, A.,
Robbins, D., Bedell, J.A. and Lakey, N.
GeneThresher methylation filtered genomic sequences from soybean
Unpublished (2003)
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fgma002d003 row: f column: 11
Seq primer: t Forward
Class: shotgun
High quality sequence stop: 633.
Location/Qualifiers
1. .633
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="fgma002d003f11"
/clone_lib="fgma filtered library (LibID: 145)"
/note="Organ: leaf; Vector: pBSK(-); Site_1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBSK(-) vector
and electroporated into E. coli cells. This is a
methylation-filtered library."

ORIGIN
Query Match 55.6%; Score 22.8; DB 8; Length 633;
Best Local Similarity 79.4%; Pred. No. 3.4e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 TGACTTTGTCCGTAATTCATAAGATCTAATTG 41
||||| ||||||| ||||||| |||||

Db 130 TGACTTTGTCCGGAACCTGCATCATCCCAATGTG 163

RESULT 10
BP678983/c
LOCUS BP678983 Osada Taira anterior neuroectoderm (ANE) PCS105 cDNA
DEFINITION library Xenopus laevis cDNA clone XL430h22ex 5', mRNA sequence.

ACCESSION BP678983
VERSION BP678983.1 GI:46026938
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 711)
AUTHORS Osada,S., Kitayama,A., Ueno,N. and Taira,M.
TITLE Expression analysis of genes which are expressed in the anterior neuroectoderm of Xenopus embryos
JOURNAL Unpublished (2004)
COMMENT Contact: Masanori Taira
Department of Biological Sciences
Graduate School of Science, University of Tokyo; CREST, Japan
Science and Technology Corporation, Japan
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-03-5841-4434
Fax: 81-03-5841-4434
Email: m.taira@biol.s.u-tokyo.ac.jp,
URL: <http://www.shigen.nig.ac.jp/nbrp/xenopus/est/>.
FEATURES
source
1..711
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL430h22ex"
/tissue_type="anterior neuroectoderm"
/dev_stage="late gastrula (stage 12.5)"
/clone_lib="Osada Taira anterior neuroectoderm (ANE)
PCSI05 cDNA library"
ORIGIN
Query Match 55.6%; Score 22.8; DB 5; Length 711;
Best Local Similarity 79.4%; Pred. No. 3.4e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CTCTCCTGACTTTGTCGTAATAATTGATAAGATCT 35
||||| ||| ||||| ||||| ||||| ||||| |||||
Db 581 CTCTCCACACATGTTCCGAAATCGATAAGATCT 548
RESULT 11
CL106907
LOCUS CL106907 1123 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-4711 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-4711,
genomic survey sequence.
ACCESSION CL106907
VERSION CL106907.1 GI:40600542
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1123)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCATTATAGG
Class: BAC ends
High quality sequence start: 14
High quality sequence stop: 759.
Location/Qualifiers
1..1123
/organism="Xenopus tropicalis"

/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-4711"
/clone_lib="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 55.6%; Score 22.8; DB 9; Length 1123;
Best Local Similarity 79.4%; Pred. No. 3.6e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 TCTCCTGACTTTGTCGTAATAATTGATAAGATCTA 36
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 471 TCTACTGACTTGCACCTTCAATTTATAAGACCTA 504
RESULT 12
CL108627/c
LOCUS CL108627 1144 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-50C23_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-50C23,
genomic survey sequence.
ACCESSION CL108627
VERSION CL108627.1 GI:40602262
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1144)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 107
High quality sequence stop: 633.
Location/Qualifiers
1..1144
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-50C23"
/clone_lib="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 55.6%; Score 22.8; DB 9; Length 1144;
Best Local Similarity 79.4%; Pred. No. 3.6e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 TCTCCTGACTTTGTCGTAATAATTGATAAGATCTA 36
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 383 TCTACTGACTTGCACCTTCAATTTATAAGACCTA 350
RESULT 13
AZ771688
LOCUS AZ771688 600 bp DNA linear GSS 16-FEB-2001
DEFINITION 1M0574H06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0574H06 F, genomic survey sequence.
ACCESSION AZ771688
VERSION AZ771688.1 GI:12894212
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 600)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0574 row: H column: 06
Seq primer: CGTTGTAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 600.
FEATURES Location/Qualifiers
source 1..600
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0574H06"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
ORIGIN
Query Match 55.1%; Score 22.6; DB 8; Length 600;
Best Local Similarity 75.7%; Pred. No. 4e+02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5 TCCTGACTTTGCCGTAATTGATAAGATCTAATTG 41
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 546 TCTACACTTTGTCAGTAAATTGAGTCAATATAATTG 582
RESULT 14
BH181335/c
LOCUS BH181335 630 bp DNA linear GSS 19-OCT-2001
DEFINITION 018_M12-21 SmbAC1 Schistosoma mansoni genomic clone 018M12 5', genomic survey sequence.
ACCESSION BH181335
VERSION BH181335.1 GI:16284256
KEYWORDS GSS.
SOURCE Schistosoma mansoni

ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 630)
AUTHORS Le Paslier,M.C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W., Williams,D.L., Johnston,D., Loverde,P.T. and Le Paslier,D.
TITLE Construction and characterization of a Schistosoma mansoni bacterial artificial chromosome library
JOURNAL Genomics 65 (2), 87-94 (2000)
MEDLINE 20247247
PUBMED 10783255
COMMENT Other GSSs: 018_M12-rev
Contact: Pierce RJ
INSERM U 167
Institut Pasteur de Lille
1 rue du Professeur A. Calmette, 59019-Lille, France
Tel: (33) (0)3 20877783
Fax: (33) (0)3 20877888
Email: Raymond.Pierce@pasteur-lille.fr
CNS sequencing ID=DG0AA018BG06CP1
Plate: 018 row: M column: 12
Seq primer: M13 -21 primer
Class: BAC ends
High quality sequence stop: 630.
FEATURES Location/Qualifiers
source 1..630
/organism="Schistosoma mansoni"
/mol_type="genomic DNA"
/strain="Puerto-Rican"
/db_xref="taxon:6183"
/clone="018M12"
/sex="mixed"
/dev_stage="cercariae"
/lab_host="Biomphalaria glabrata"
/clone_lib="SmbAC1"
/note="Vector: pBelobAC 11; Site 1: Hind III; Partially Hind III digested and size-selected S. mansoni cercarial DNA was ligated into Hind III digested pBelobAC 11 vector and used to transform E. coli DH10B. The complete library contains 23808 clones from 4 independent sizing-ligation-transformations. Average insert size ranges from 70-127 kb and genome coverage is 7.9-fold."
ORIGIN
Query Match 55.1%; Score 22.6; DB 8; Length 630;
Best Local Similarity 86.2%; Pred. No. 4e+02;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 TCCTGACTTTGCCGTAATTGATAAGAT 33
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 TCCAGAAATTGCTGTAAATTGATAATAT 264
RESULT 15
CNS07MZF/c
LOCUS CNS07MZF 630 bp DNA linear GSS 30-NOV-2001
DEFINITION T7 end of clone 018BG06 of library SmbAC1 from strain Puerto-Rican of Schistosoma mansoni, genomic survey sequence.
ACCESSION AL618285
VERSION AL618285.1 GI:16031789
KEYWORDS GSS.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 630)
AUTHORS Le Paslier,M.C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W., Williams,D.L., Johnston,D., Loverde,P.T. and Le Paslier,D.
TITLE Construction and characterization of a Schistosoma mansoni bacterial artificial chromosome library
JOURNAL Genomics 65 (2), 87-94 (2000)
MEDLINE 20247247
PUBMED 10783255

REFERENCE 2 (bases 1 to 630)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segrei@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Partially Hind III digested and size-selected S. mansoni cercarial
DNA was ligated into Hind III digested pBelobAC 11 vector and used
to transform E. coli DH10B. The complete library contains 23808
clones from 4 independent sizing-ligation-transformations. Average
insert size ranges from 70-127 kb and genome coverage is 7.9-fold.
FEATURES
Location/Qualifiers
1..630
/organism="Schistosoma mansoni"
/mol_type="genomic DNA"
/strain="Puerto-Rican"
/db_xref="taxon:6183"
/clone="018BG06"
/clone_lib="SmbAC1"
/note="end : T7"
ORIGIN
Query Match 55.1%; Score 22.6; DB 9; Length 630;
Best Local Similarity 86.2%; Pred. No. 4e+02;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 TCCTGACTTTGTCCGTAATTCGATAAGAT 33
||| ||||| ||||| ||||| ||||| |||||
Db 292 TCCAGAAATTGTCTGTAAATTGATAATAT 264

Search completed: November 3, 2004, 09:43:29
Job time : 711.123 secs

This Page Blank (uspto)

SECRET

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:44:23 ; Search time 83.8636 Seconds
(without alignments)
2566.384 Million cell updates/sec

Title: US-09-933-267A-1_COPY_64311_64351

Perfect score: 41

Sequence: 1 tctctcctgactttgtccgt.....aattgataagatctaatttg 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	110000	6	ABA90193_0
2	41	100.0	110000	6	ABQ87681_0
3	41	100.0	110000	8	ABX33717_0
4	39.4	96.1	110000	10	ADB81391_Partial g
5	23.8	58.0	1347	10	ADH82504
6	23.8	58.0	1368	4	AAS51295
7	23.8	58.0	1389	8	ACA18377
8	23.8	58.0	1392	4	AAS52928
9	23.4	57.1	591	4	ABA63739
10	23.4	57.1	591	4	AAI43851
11	23.4	57.1	591	4	AAK37973
12	23.4	57.1	591	4	ABS37592
13	23.4	57.1	591	6	ABS11972
14	23	56.1	1360	12	ADI16270
15	23	56.1	2069	10	ADB62733
16	22.4	54.6	658	3	AAA44181
17	22.4	54.6	2709	8	ACA51279
18	22.4	54.6	101786	3	AAF22293
19	22	53.7	32172	4	AAL03473
20	21.8	53.2	1788	8	ABX34799
21	21.8	53.2	2274	5	AAS65768

C 22	21.8	53.2	2386	5	AAS86931	Aas86931 DNA encod
C 23	21.8	53.2	2621	10	ADB62795	Adb62795 Human cDN
C 24	21.8	53.2	2978	10	ADC37641	Adc37641 Human nuc
25	21.4	52.2	1541	8	ACA49041	Aca49041 Prokaryot
26	21.4	52.2	163350	6	AAD46127	Aad46127 Human tum
C 27	21.2	51.7	16831	4	AAK73455	Aak73455 Human imm
C 28	21.2	51.7	89900	12	ADO79404	Ado79404 DPFB3 regi
C 29	21	51.2	638	4	AAH08204	Aah08204 Human cDN
30	21	51.2	819	6	ABZ32168	Abz32168 Candida a
C 31	21	51.2	2356	4	AAH16613	Aah16613 Human cDN
C 32	21	51.2	3364	11	ADM01350	Adm01350 Human cDN
C 33	21	51.2	9540	12	ADO52562	Ado52562 Human imm
C 34	21	51.2	9540	12	ADP20074	Adp20074 Human imm
C 35	21	51.2	10872	6	ABK31389	Abk31389 Signal tr
C 36	21	51.2	10872	6	ABL70336	Ab170336 Chemicall
C 37	21	51.2	10872	6	AAS61289	Aas61289 Human gen
C 38	20.8	50.7	375	5	ABV10419	Abv10419 Human pro
C 39	20.8	50.7	375	5	ABV01250	Abv01250 Human pro
C 40	20.8	50.7	409	5	ABV40557	Abv40557 Human pro
C 41	20.8	50.7	409	5	ABV31589	Abv31589 Human pro
C 42	20.8	50.7	423	8	ACF73919	Acf73919 Staphyloc
C 43	20.8	50.7	598	6	ABK53513	Abk53513 Human eos
C 44	20.8	50.7	951	10	ADF01416	Adf01416 Bacterial
C 45	20.8	50.7	1557	2	AAQ30973	Aaq30973 HIV gp160

ALIGNMENTS

RESULT 1
ABA90193_0
WP Sequence split into 5 fragments LOCUS ABA90193 Accession ABA90193
WP Fragment Name Begin End
WP ABA90193_0 1 110000
WP ABA90193_1 100001 210000
WP ABA90193_2 200001 310000
WP ABA90193_3 300001 410000
WP ABA90193_4 400001 465237
ID ABA90193 standard; DNA; 465237 BP.
XX
AC ABA90193;
XX
DT 11-FEB-2002 (first entry)
XX
DE Human oestrogen receptor alpha gene.
XX
KW Human; oestrogen receptor alpha; ESR-alpha; ER; chromosome 6; Syne-2;
KW synaptic nuclei expressed gene 2; haplotype; cytostatic; osteopathic;
KW cardiant; vasotropic; gene therapy; vaccine; cancer; osteoporosis;
KW cardiovascular disease; oestrogen receptor; ds.
XX
OS Homo sapiens.
XX
PN WO200162969-A2.
XX
PD 30-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005358.
XX
PR 22-FEB-2000; 2000US-0183756P.
PR 20-OCT-2000; 2000US-00692414.
PR 24-JAN-2001; 2001US-00768184.
XX
PA (PEKE) PE CORP NY.
XX
PI Kalush F, Cassel MJ, Hwang SS, Winn-Deen ES;
XX
DR WPI; 2002-041152/05.
DR P-PSDB; AAG68251.
XX
PT Novel variant of estrogen receptor alpha polypeptide useful for
PT determining the biological activity of a protein for high throughput
PT screening and for raising antibodies that elicit an immune response in

41 Frag

FT host.
XX
PS Example; Fig 1 page 1-93; 333pp; English.
XX
CC The present invention describes an isolated peptide (I) consisting of an
CC amino acid sequence selected from: (a) the amino acid sequence of a
CC variant of the oestrogen receptor alpha (ESR-alpha) protein in AAG68251;
CC or (b) a fragment comprising at least 10 contiguous amino acids of the
CC protein in AAG68251. (I) has cytostatic, osteopathic, cardiant and
CC vasotropic activities, and can be used in gene therapy and vaccine
CC production. (I) is useful for identifying an agent that binds to (I), by
CC contacting (I) with an agent and assaying the contacted mixture to
CC determine whether a complex is formed with the agent bound to the
CC peptide. A polynucleotide (II), encoding (I), is useful in the
CC development of diagnostics and therapies for diseases and disorders
CC mediated/modulated by an oestrogen receptor (ER). (II) is also useful in
CC gene therapy for treating cancer, osteoporosis and cardiovascular
CC diseases. The human ESR-alpha gene is located on chromosome 6. The
CC present sequence represents the human ESR-alpha gene, which is given in
CC the exemplification of the present invention
XX
SQ Sequence 465237 BP; 133988 A; 89578 C; 93946 G; 147721 T; 0 U; 4 Other;

Query Match 100.0%; Score 41; DB 6; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCGCGAAATTGATAAGATCTAATTG 41
DB 64311 TCTCTCCTGACTTTGTCGCGAAATTGATAAGATCTAATTG 64351

RESULT 2
ABQ87681_0
WP Sequence split into 5 fragments LOCUS ABQ87681 Accession Abq87681
WP Fragment Name Begin End
WP ABQ87681_0 1 110000
WP ABQ87681_1 100001 210000
WP ABQ87681_2 200001 310000
WP ABQ87681_3 300001 410000
WP ABQ87681_4 400001 465237
ID ABQ87681 standard; DNA; 465237 BP.
XX
AC ABQ87681;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human oestrogen receptor alpha gene.
XX
KW Human; oestrogen; receptor; oestrogen receptor alpha; cytostatic;
KW osteopathic; cardiant; cancer; osteoporosis; cardiovascular disorder;
KW chromosome 6q25.1; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(18783,T)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18937,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT exon 18941..19032
FT /*tag= b
FT /number= 1
FT intron 19033..52817
FT /*tag= b
FT /number= 1
FT /cons_splice= (5'site:YES,3'site:NO)
FT variation replace(19034,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT exon 52818..52940

FT /*tag= b
FT /number= 2
FT replace(52877,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(52901,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 52941..64149
FT /*tag= b
FT /number= 2
FT /cons_splice= (5'site:YES,3'site:NO)
FT 64150..64280
FT /*tag= b
FT /number= 3
FT 64281..166227
FT /*tag= b
FT /number= 3
FT /cons_splice= (5'site:YES,3'site:NO)
FT replace(64331,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 166228..166322
FT /*tag= b
FT /number= 4
FT 166323..168001
FT /*tag= b
FT /number= 4
FT replace(167950,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(167989,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 168002..168120
FT /*tag= b
FT /number= 5
FT replace(168054,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 168121..169542
FT /*tag= b
FT /number= 5
FT /cons_splice= (5'site:YES,3'site:NO)
FT 169543..169825
FT /*tag= b
FT /number= 6
FT replace(169812,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(169823,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 169826..169866
FT /*tag= b
FT /number= 6
FT /cons_splice= (5'site:YES,3'site:NO)
FT 169867..170678
FT /*tag= b
FT /number= 7
FT replace(170035,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(170068,T)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(170256,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(170368,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(170487,C)
FT variation

FT / *tag= a /standard_name= "Single nucleotide polymorphism"
FT 170679. .204911
FT / *tag= b
FT /number= 7
FT 204912. .205102
FT / *tag= b
FT /number= 8
FT 205103. .242969
FT / *tag= b
FT /number= 8
FT 242970. .243086
FT / *tag= b
FT /number= 9
FT replace(243055,T)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 243087. .306167
FT / *tag= b
FT /number= 9
FT replace(243187,C)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 306168. .306503
FT / *tag= b
FT /number= 10
FT replace(306292,A)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(306382,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 306504. .373639
FT / *tag= b
FT /number= 10
FT 373640. .373778
FT / *tag= b
FT /number= 11
FT 373779. .422963
FT / *tag= b
FT /number= 11
FT 422964. .423097
FT / *tag= b
FT /number= 12
FT replace(423067,C)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 423098. .456353
FT / *tag= b
FT /number= 12
FT replace(423149,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(423163,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(423220,A)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(423232,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(423258,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT 456354. .456537
FT / *tag= b
FT /number= 13
FT 456538. .460700
FT / *tag= b
FT /number= 13
FT replace(459706,C)
FT / *tag= a

FT variation /standard_name= "Single nucleotide polymorphism"
FT replace(459832,A)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(459913,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(460024,G)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(460056,T)
FT / *tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(460159,G)
FT / *tag= a

Query Match 100.0%; Score 41; DB 6; Length 110000;
Best Local Similarity 100.0%; Pred. NO. 5.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAAATTGATAAGATCTAATTG 41
|||
Db 64311 TCTCTCCTGACTTTGTCCGTAAATTGATAAGATCTAATTG 64351

RESULT 3
ABX33717_0
WP Sequence split into 5 fragments LOCUS ABX33717 Accession Abx33717
WP Fragment Name Begin End
WP ABX33717_0 1 110000
WP ABX33717_1 100001 210000
WP ABX33717_2 200001 310000
WP ABX33717_3 300001 410000
WP ABX33717_4 400001 465237
ID ABX33717 standard; DNA; 465237 BP.

ABX33717;

26-FEB-2003 (first entry)

Gene encoding human oestrogen receptor alpha protein (ESR1).

Human; oestrogen receptor alpha; ESR1; cancer; osteoporosis;
cardiovascular disorder; variant oestrogen receptor; ESR1 haplotype;
ESR1 polymorphism detection; cytostatic; osteopathic; cardiant;
chromosome 6q25.1; gene; single nucleotide polymorphism; SNP; ds.

Homo sapiens.

PH Key Location/Qualifiers
FT variation replace(18783,T)
FT / *tag= ao
FT /standard_name= "SNP"
FT variation replace(18937,C)
FT / *tag= ap
FT /standard_name= "SNP"
FT 18941. .19032
FT / *tag= a
FT /number= 1G
FT 19033. .52817
FT / *tag= b
FT /number= 1G
FT variation replace(19034,C)
FT / *tag= aq
FT /standard_name= "SNP"
FT 52818. .52940
FT / *tag= c
FT /number= 1F
FT variation replace(52877,A)
FT / *tag= an
FT /standard_name= "SNP"
FT variation replace(52901,A)
FT / *tag= am

FT	intron	/standard_name= "SNP"
FT		52941. .64149
FT		/*tag= d
FT	exon	/numBer= 1F
FT		64150. .64280
FT		/*tag= e
FT	intron	/number= 1E
FT		64281. .166227
FT		/*tag= f
FT	variation	/number= 1E
FT		replace(64331,G)
FT		/*tag= al
FT	exon	/standard_name= "SNp"
FT		166228. .166322
FT		/*tag= g
FT	intron	/number= 1D
FT		166323. .168001
FT		/*tag= h
FT	variation	/number= 1D
FT		replace(167950,G)
FT		/*tag= ai
FT	variation	/standard_name= "SNP"
FT		replace(167989,G)
FT		/*tag= aj
FT	exon	/standard_name= "SNP"
FT		168002. .168120
FT		/*tag= i
FT	variation	/number= 1C
FT		replace(168054,G)
FT		/*tag= ak
FT	intron	/standard_name= "SNP"
FT		168121. .169542
FT		/*tag= j
FT	exon	/number= 1C
FT		169543. .169825
FT		/*tag= k
FT	variation	/number= 1B
FT		replace(169812,G)
FT		/*tag= ag
FT	variation	/standard_name= "SNP"
FT		replace(169823,G)
FT		/*tag= ah
FT	intron	/standard_name= "SNP"
FT		169826. .169866
FT		/*tag= l
FT	exon	/number= 1B
FT		169867. .170678
FT		/*tag= m
FT	variation	/number= 1A
FT		replace(170035,A)
FT		/*tag= ab
FT	variation	/standard_name= "SNP"
FT		replace(170068,T)
FT		/*tag= ac
FT	variation	/standard_name= "SNp"
FT		replace(170256,C)
FT		/*tag= ad
FT	variation	/standard_name= "SNP"
FT		replace(170368,G)
FT		/*tag= ae
FT	variation	/standard_name= "SNP"
FT		replace(170487,C)
FT		/*tag= af
FT	intron	/standard_name= "SNP"
FT		170679. .204911
FT		/*tag= n
FT	exon	/number= 1A
FT		204912. .205102
FT		/*tag= o
FT		/number= 2
FT	intron	205103. .242969
FT		/*tag= P
FT		/number= 2

FT	exon	242970. .243086
FT		/*tag= q
FT	variation	/number= 3
FT		replace(243055,T)
FT		/*tag= as
FT	intron	/standard_name= "SNP"
FT		243087. .306167
FT		/*tag= r
FT	variation	/number= 3
FT		replace(243187,C)
FT		/*tag= ar
FT	exon	/standard_name= "SNP"
FT		306168. .306503
FT		/*tag= s
FT	variation	/number= 4
FT		replace(306292,A)
FT		/*tag= at
FT	variation	/standard_name= "SNP"
FT		replace(306382,G)
FT		/*tag= au
FT	intron	/standard_name= "SNP"
FT		306504. .373639
FT		/*tag= t
FT	exon	/number= 4
FT		373640. .373778
FT		/*tag= u
FT	intron	/number= 5
FT		373779. .422963
FT		/*tag= v
FT	exon	/number= 5
FT		422964. .423097
FT		/*tag= w
FT	variation	/number= 6
FT		replace(423067,C)
FT		/*tag= av
FT	intron	/standard_name= "SNP"
FT		423098. .456353
FT		/*tag= x
FT	variation	/number= 6
FT		replace(423149,G)
FT		/*tag= aw
FT	variation	/standard_name= "SNP"
FT		replace(423163,G)
FT		/*tag= ax
FT	variation	/standard_name= "SNP"
FT		replace(423220,A)
FT		/*tag= ay
FT	variation	/standard_name= "SNP"
FT		replace(423232,G)
FT		/*tag= az
FT	variation	/standard_name= "SNP"
FT		replace(423258,G)
FT		/*tag= ba
FT	exon	/standard_name= "SNP"
FT		456354. .456537
FT		/*tag= y
FT	intron	/number= 7
FT		456538. .460700
FT		/*tag= z
FT	variation	/number= 7
FT		replace(459706,C)
FT		/*tag= bb
FT	variation	/standard_name= "SNP"
FT		replace(459832,A)
FT		/*tag= bc
FT	variation	/standard_name= "SNP"
FT		replace(459913,G)
FT		/*tag= bd
FT	variation	/standard_name= "SNP"
FT		replace(460024,G)
FT		/*tag= be
FT	variation	/standard_name= "SNP"
FT		replace(460056,T)

FT /*tag= bf
FT /standard_name= "SNP"
FT replace(460159,G)
FT /*tag= bg
FT /standard_name= "SNP"
FT replace(460553,C)
FT /*tag= bh
FT /standard_name= "SNP"

Query Match 100.0%; Score 41; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
 |||||
Db 64311 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 64351

RESULT 4
ADB81391_0
WP Sequence split into 4 fragments LOCUS ADB81391 Accession Adb81391
WP Fragment Name Begin End
WP ADB81391_0 1 110000
WP ADB81391_1 100001 210000
WP ADB81391_2 200001 310000
WP ADB81391_3 300001 392000
ID ADB81391 standard; DNA; 392000 BP.
XX
AC ADB81391;
XX
DT 04-DEC-2003 (first entry)
XX
DE Partial genomic DNA sequence of the human oestrogen receptor alpha mRNA.
XX
KW gene; human; ds; oestrogen receptor alpha; ESR-alpha;
KW oestrogen receptor 1; ESR1; NR3A1; bone maintenance;
KW cardiovascular system; cancer; gene therapy; hyperproliferative disease;
KW inflammation; tumour formation; infection; cytostatic; antiinflammatory;
KW antimicrobial.
XX
OS Homo sapiens.
XX
PN WO2003052072-A2.
XX
PD 26-JUN-2003.
XX
PF 13-DEC-2002; 2002WO-US040083.
XX
PR 18-DEC-2001; 2001US-00027983
XX (ISIS-) ISIS PHARM INC.
PA
PI Dobie KW, Roach MP;
XX
DR WPI; 2003-577322/54.
XX
PT New antisense compound targeted to nucleic acid encoding estrogen
PT receptor alpha and inhibiting expression of estrogen receptor alpha,
PT useful for treating a disease or condition e.g. a hyperproliferative
PT disease.
XX
PS Example 15; Page 99-210; 232pp; English.
XX
CC This invention relates to human oestrogen receptor alpha (ESR-alpha), and
CC the novel antisense oligonucleotides that modulate its expression. The
CC oestrogen receptor alpha protein is also known as oestrogen receptor 1,
CC ESR1, and NR3A1. Oestrogen, the steroid hormone ligand of ESR-alpha, is
CC important for bone maintenance and plays a protective role in the
CC cardiovascular system, as well as being required for normal sexual
CC maturation through promoting growth and differentiation. Splice variants
CC of ESR-alpha, however, have been associated with various cancers
CC including the breast and pituitary. Accordingly, antisense
CC oligonucleotides that inhibit the expression of ESR-alpha in cells or

CC tissues can be used in gene therapy to treat conditions such as
CC hyperproliferative disease, inflammation, tumour formation and to prevent
CC or delay infection. As such, the present invention describes these
CC antisense oligos as having cytostatic, antiinflammatory and antimicrobial
CC activities. This polynucleotide is the partial genomic sequence of the
CC human oestrogen receptor alpha DNA of the invention.
XX
SQ Sequence 392000 BP; 113320A; 73484C; 75996G; 117294T; 0U; 119060Other;

Query Match 96.1%; Score 39.4; DB 10; Length 110000;
Best Local Similarity 97.6%; Pred. No. 2.2e-05;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
 |||||
Db 12773 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 12813

RESULT 5
ADH82504/c
ID ADH82504 standard; DNA; 1347 BP.
XX
AC ADH82504;
XX
DT 22-APR-2004 (first entry)
XX
DE Enterococcus faecalis polynucleotide #389.
XX
KW Enterococcus faecalis infection; transcription regulatory element;
KW antibacterial; gene; ds.
XX
OS Enterococcus faecalis.
XX
PN US6617156-B1.
XX
PD 09-SEP-2003.
XX
PF 13-AUG-1998; 98US-00134000.
XX
PR 15-AUG-1997; 97US-0055778P.
XX
PA (DOUC/) DOUCETTE-STAMM L A.
PA (BUSH/) BUSH D.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2003-895394/82.
DR P-PSDB; ADH85909.
XX
PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX
PS Disclosure; SEQ ID NO 389; 193pp; English.
XX
CC The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polynucleotide of the invention.
XX
SQ Sequence 1347 BP; 322 A; 217 C; 328 G; 480 T; 0 U; 0 Other;

Query Match 58.0%; Score 23.8; DB 10; Length 1347;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TCTCCTGACTTTGTCGFAAATTGATAAGATCTAA 37
1151 TCTCCTTTTCATTGTCGTAATTGATAATTCCTAA 1117

RESULT 6
AAS51295/c
ID AAS51295 standard; DNA; 1368 BP.
AC AAS51295;
DT 13-FEB-2002 (first entry)
DE Enterococcus faecalis DNA for cellular proliferation protein #72.
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX Enterococcus faecalis.
OS WO200170955-A2.
PN 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
PF 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR P-PSDB; AAU33436.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.

Claim 27; SEQ ID NO 3877; 511pp; English.
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence encodes an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1368 BP; 330 A; 220 C; 332 G; 486 T; 0 U; 0 Other;
SQ

Query Match 58.0%; Score 23.8; DB 4; Length 1368;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

3 TCTCCTGACTTTGTCGFAAATTGATAAGATCTAA 37

Db 1193 TCTCCTTTTCATTGTCGTAATTGATAATTCCTAA 1159

RESULT 7
ACA18377/c
ID ACA18377 standard; DNA; 1389 BP.
XX ACA18377;
AC ACA18377;
DT 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #34.
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
KW Enterococcus faecalis.
OS WO200277183-A2.
PN 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
PF 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU14507.

New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 6247; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than S. aureus, S. typhimurium,
K. pneumoniae or P. aeruginosa. The present sequence is one of the target
prokaryotic essential genes. Note: The sequence data for this patent did

CC not form part of the printed specification; but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1389 BP; 337 A; 224 C; 337 G; 491 T; 0 U; 0 Other;
Query Match 58.0%; Score 23.8; DB 8; Length 1389;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 TCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAA 37
||||| | ||||| ||||| ||||| |||||
Db 1196 TCTCCTTCATTGTCCGTAATAATTGATAATTCCTAA 1162

RESULT 8
AAS52928/c
ID AAS52928 standard; DNA; 1392 BP.
XX
AC AAS52928;
XX
DT 13-FEB-2002 (first entry)
XX
DE Enterococcus faecalis DNA for cellular proliferation protein #356.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU35069.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 6565; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1392 BP; 341 A; 223 C; 337 G; 491 T; 0 U; 0 Other;
Query Match 58.0%; Score 23.8; DB 4; Length 1392;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 TCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAA 37
||||| | ||||| ||||| ||||| |||||
Db 1196 TCTCCTTCATTGTCCGTAATAATTGATAATTCCTAA 1162

RESULT 9
ABA63739
ID ABA63739 standard; DNA; 591 BP.
XX
AC ABA63739;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #12044.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
PS Claim 1; SEQ ID NO 12044; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 591 BP; 188 A; 89 C; 98 G; 216 T; 0 U; 0 Other;
Query Match 57.1%; Score 23.4; DB 4; Length 591;
Best Local Similarity 73.2%; Pred. No. 21;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAAATTG 41
||||| | ||||| ||||| ||||| |||||
Db 80 TCTCTACTGACTTTCTCTTAATTAGCTTAGTGTAAATTG 120

```
RESULT 10
AAI43851
ID AAI43851 standard; DNA; 591 BP.
XX
AC AAI43851;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #12537 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US0000663.
PR
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-48897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 12537; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 591 BP; 188 A; 89 C; 98 G; 216 T; 0 U; 0 Other;

Query Match 57.1%; Score 23.4; DB 4; Length 591;
Best Local Similarity 73.2%; Pred. No. 21;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTCTCTGACTTTGTCGTAATTGATAAGATCTAATTG 41
Db ||||| ||||| || ||| ||| |||||
80 TCTCTACTGACTTCTCTTTAATTAGCTTAGTGTAAATTG 120

RESULT 11
AAK37973
ID AAK37973 standard; DNA; 591 BP.
XX
AC AAK37973;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 12530.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
```

```
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 12530; 658pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 591 BP; 188 A; 89 C; 98 G; 216 T; 0 U; 0 Other;

Query Match 57.1%; Score 23.4; DB 4; Length 591;
Best Local Similarity 73.2%; Pred. No. 21;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTCTCTGACTTTGTCGTAATTGATAAGATCTAATTG 41
Db ||||| ||||| || ||| ||| |||||
80 TCTCTACTGACTTCTCTTTAATTAGCTTAGTGTAAATTG 120

RESULT 12
ABS37592
ID ABS37592 standard; DNA; 591 BP.
XX
AC ABS37592;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 12582.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
```


XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX Claim 1; SEQ ID NO 12582; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 591 BP; 188 A; 89 C; 98 G; 216 T; 0 U; 0 Other;
SQ Query Match 57.1%; Score 23.4; DB 4; Length 591;
Best Local Similarity 73.2%; Pred. No. 21;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TCTCTCTGACTTTGTCGTAATTAAGATCTAATTG 41
Db 80 TCTCTACTGACTTCTCTTAATTAGCTTAGTGTAAATTG 120
RESULT 13
ABS11972
ID ABS11972 standard; DNA; 591 BP.
AC ABS11972;
XX 19-AUG-2002 (first entry)
DE Human genome-derived single exon probe from lung SEQ ID No 11963.
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
OS WO200186003-A2.
PN 15-NOV-2001.
XX 30-JAN-2001; 2001WO-US000665.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX Claim 1; SEQ ID NO 11963; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation of each of the exons in several
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 591 BP; 188 A; 89 C; 98 G; 216 T; 0 U; 0 Other;
SQ Query Match 57.1%; Score 23.4; DB 6; Length 591;
Best Local Similarity 73.2%; Pred. No. 21;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TCTCTCTGACTTTGTCGTAATTAAGATCTAATTG 41
Db 80 TCTCTACTGACTTCTCTTAATTAGCTTAGTGTAAATTG 120
RESULT 14
ADI16270/c
ID ADI16270 standard; DNA; 1360 BP.
XX AC ADI16270;
XX 22-APR-2004 (first entry)
DT Human nucleic acid-associated protein (NAAP) coding sequence #5.
DE human; nucleic acid-associated protein; NAAP; autoimmune disorder;
XX inflammatory disorder; AIDS; allergy; infection; metabolic disorder;
KW obesity; reproductive disorder; infertility; neurological disorder;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 06:51:04 ; Search time 74.6435 Seconds
(without alignments)
2816.475 Million cell updates/sec

Title: US-09-933-267A-1_COPY_64311_64351

Perfect score: 41
Sequence: 1 tctctcctgactttgtcgt.....aattgataagatctaatttg 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	41	100.0	465237	9	US-09-933-267A-1
2	39.4	96.1	392000	15	US-10-027-983-11
3	39.4	96.1	392000	15	US-10-448-753-11
4	24	58.5	149062	17	US-10-367-094-93
5	23.8	58.0	1368	9	US-09-815-242-3877
6	23.8	58.0	1389	16	US-10-282-122A-6247
7	23.8	58.0	1392	9	US-09-815-242-6565
8	23.4	57.1	591	9	US-09-864-761-14894
9	23	56.1	2069	15	US-10-104-047-887
10	23	56.1	78072	15	US-10-085-117-154
11	22.4	54.6	2709	16	US-10-282-122A-39149
12	22	53.7	5865	15	US-10-369-493-46473

13	22	53.7	32172	10	US-09-764-891-6161	Sequence 6161, Ap
14	21.8	53.2	434	16	US-10-424-599-23624	Sequence 23624, A
15	21.8	53.2	2047	16	US-10-425-114-12000	Sequence 12000, A
16	21.8	53.2	2070	16	US-10-424-599-121294	Sequence 121294, A
17	21.8	53.2	2621	15	US-10-104-047-949	Sequence 949, App
18	21.6	52.7	982	16	US-10-424-599-100224	Sequence 100224, A
19	21.6	52.7	2012	9	US-09-887-576-242	Sequence 242, App
20	21.4	52.2	1541	16	US-10-282-122A-36911	Sequence 36911, A
21	21.4	52.2	3225	17	US-10-437-963-24292	Sequence 24292, A
22	21.2	51.7	474	18	US-10-674-124A-11922	Sequence 11922, A
23	21.2	51.7	824	16	US-10-424-599-41879	Sequence 41879, A
24	21	51.2	437	13	US-10-027-632-182271	Sequence 182271, A
25	21	51.2	437	15	US-10-027-632-182271	Sequence 182271, A
26	21	51.2	552	13	US-10-027-632-282130	Sequence 282130, A
27	21	51.2	552	13	US-10-027-632-282131	Sequence 282131, A
28	21	51.2	552	13	US-10-027-632-282132	Sequence 282132, A
29	21	51.2	552	13	US-10-027-632-282133	Sequence 282133, A
30	21	51.2	552	15	US-10-027-632-282130	Sequence 282130, A
31	21	51.2	552	15	US-10-027-632-282131	Sequence 282131, A
32	21	51.2	552	15	US-10-027-632-282132	Sequence 282132, A
33	21	51.2	552	15	US-10-027-632-282133	Sequence 282133, A
34	21	51.2	819	15	US-10-032-585-6455	Sequence 6455, Ap
35	21	51.2	3364	16	US-10-108-260A-35	Sequence 35, Appl
36	21	51.2	9540	17	US-10-325-468-46	Sequence 46, Appl
37	21	51.2	10872	16	US-10-221-613-250	Sequence 250, App
38	20.8	50.7	455	17	US-10-437-963-72997	Sequence 72997, A
39	20.8	50.7	2238	15	US-10-032-585-6125	Sequence 6125, Ap
40	20.8	50.7	2370	16	US-10-282-122A-18027	Sequence 18027, A
41	20.8	50.7	2698	16	US-10-469-199-5	Sequence 5, Appli
42	20.8	50.7	2709	16	US-10-282-122A-39991	Sequence 39991, A
43	20.8	50.7	8339	8	US-08-781-986A-175	Sequence 175, App
44	20.8	50.7	8339	16	US-10-329-624-175	Sequence 175, App
45	20.8	50.7	13069	9	US-09-764-869-1850	Sequence 1850, Ap

ALIGNMENTS

RESULT 1

US-09-933-267A-1
; Sequence 1, Application US/099333267A
; Patent No. US20020123095A1
; GENERAL INFORMATION:
; APPLICANT: Kalush, Francis et al.
; TITLE OF INVENTION: Estrogen receptor alpha variants and methods of detection thereof
; FILE REFERENCE: CL000258C14
; CURRENT APPLICATION NUMBER: US/09/933,267A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/160626
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 60/183756
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/692414
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/768184
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09/804076
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/826314
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 465237
; TYPE: DNA
; ORGANISM: human
US-09-933-267A-1

Query Match 100.0%; Score 41; DB 9; Length 465237;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
|||||
Db 64311 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 64351

RESULT 2

US-10-027-983-11
; Sequence 11, Application US/10027983
; Publication No. US20030139360A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137740
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 137742
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (138122)...(138221)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 145507
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 151967
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (151967)...(1542066)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 154217
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (164037)...(164136)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (174657)...(174756)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (186224)...(186323)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (195242)...(195341)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 202703
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (202771)...(202870)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (206246)...(215602)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (218126)...(218225)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (220360)...(220459)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (222717)...(222816)
; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc feature
; LOCATION: (223981)...(224080)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (227487)...(227586)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (230157)...(230256)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (232299)...(232398)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (236552)...(2366651)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (238789)...(248788)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: exon
; LOCATION: (118288)...(119101)
; OTHER INFORMATION: exon 1C
; NAME/KEY: exon:intron junction
; LOCATION: (151129)...(151130)
; OTHER INFORMATION: exon 5:intron 5
; NAME/KEY: exon:intron junction
; LOCATION: (299248)...(299249)
; OTHER INFORMATION: exon 9:intron 9
; NAME/KEY: exon:intron junction
; LOCATION: (348578)...(348579)
; OTHER INFORMATION: exon 10:intron 10
; NAME/KEY: intron
; LOCATION: (348579)...(381838)
; OTHER INFORMATION: intron 10
; NAME/KEY: intron:exon junction
; LOCATION: (386185)...(386186)
; OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Query Match 96.1%; Score 39.4; DB 15; Length 392000;
Best Local Similarity 97.6%; Pred. No. 0.00011;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
|||||
Db 12773 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 12813

RESULT 3

US-10-448-753-11
; Sequence 11, Application US/10448753
; Publication No. US20030211611A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/448,753
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US/10/027,983
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137740
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137742
; OTHER INFORMATION: unknown

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (138122)...(138221)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: 145507
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 151967
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (151967)...(1542066)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: 154217
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (164037)...(164136)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (174657)...(174756)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (186224)...(186323)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (195242)...(195341)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: 202703
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (202771)...(202870)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (206246)...(215602)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (218126)...(218225)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (220360)...(220459)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (222717)...(222816)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (223981)...(224080)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (227487)...(227586)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (230157)...(230256)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (232299)...(232398)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (236552)...(2366651)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (238789)...(248788)
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: exon
LOCATION: (118288)...(119101)
OTHER INFORMATION: exon 1C
FEATURE:
NAME/KEY: exon:intron junction
LOCATION: (151129)...(151130)
OTHER INFORMATION: exon 5:intron 5
FEATURE:
NAME/KEY: exon:intron junction
LOCATION: (299248)...(299249)
OTHER INFORMATION: exon 9:intron 9
FEATURE:
NAME/KEY: exon:intron junction
LOCATION: (348578)...(348579)
OTHER INFORMATION: exon 10:intron 10
FEATURE:
NAME/KEY: intron
LOCATION: (348579)...(381838)
OTHER INFORMATION: intron 10
FEATURE:
NAME/KEY: intron:exon junction
LOCATION: (386185)...(386186)
OTHER INFORMATION: intron 11:exon 12
US-10-448-753-11

Query Match      96.1%; Score 39.4; DB 15; Length 392000;
Best Local Similarity 97.6%; Pred. No. 0.00011;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 TCTCTCCTGACTTTGTCCGTAATGTGATAAGATCTAATTG 41
Db      12773 TCTCTCCTGACTTTGTCCGTAATGTGATAAGATCTAATTG 12813

RESULT 4
US-10-367-094-93
; Sequence 93, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 149062
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-367-094-93

Query Match      58.5%; Score 24; DB 17; Length 149062;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy      1 TCTCTCCTGACTTTGTCCGTAATGTGATAAGATCTAATT 40
Db      124343 TCACTGATAACTTACTCAGTGATATATAAGATCTAATT 124382
```

RESULT 5

```
US-09-815-242-3877/c
; Sequence 3877, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3877
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-815-242-3877
```

```
Query Match      58.0%; Score 23.8; DB 9; Length 1368;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      3  TCTCCTGACTTTGTCCGTAATGATAAGATCTAA 37
          |||||  |||||  |||||  |||||  |||||  |||||
Db      1193 TCTCCTTTCATTGTCCGTAATTGATAAATTCCTAA 1159
```

RESULT 6

```
US-10-282-122A-6247/c
; Sequence 6247, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6247
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-282-122A-6247
```

```
Query Match      58.0%; Score 23.8; DB 16; Length 1389;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      3  TCTCCTGACTTTGTCCGTAATGATAAGATCTAA 37
          |||||  |||||  |||||  |||||  |||||  |||||
Db      1196 TCTCCTTTCATTGTCCGTAATTGATAAATTCCTAA 1162
```

RESULT 7

```
US-09-815-242-6565/c
; Sequence 6565, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6565
; LENGTH: 1392
; TYPE: DNA
```


; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1392)
US-09-815-242-6565

Query Match 58.0%; Score 23.8; DB 9; Length 1392;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TCTCCTGACTTTGTCGTAATAATTGATAAGATCTAA 37
||||| | ||||| ||||| ||||| ||||| |||||
Db 1196 TCTCCTTTTCATTGTCGTAATAATTGATAATTCCTAA 1162

RESULT 8
US-09-864-761-14894
; Sequence 14894, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14894
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007719.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
US-09-864-761-14894

Query Match 57.1%; Score 23.4; DB 9; Length 591;
Best Local Similarity 73.2%; Pred. No. 40;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAAATTG 41
||||| | ||||| ||||| ||||| ||||| |||||
Db 80 TCTCTACTGACTTTCTCTTAAATTAGCTTAGTGTAAATTG 120

RESULT 9
US-10-104-047-887/c
; Sequence 887, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 887
; LENGTH: 2069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-887

Query Match 56.1%; Score 23; DB 15; Length 2069;
Best Local Similarity 74.4%; Pred. No. 79;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAAATT 39
||||| | ||||| ||||| ||||| ||||| |||||
Db 855 TCTCTCATGTCTTCGATAGGAAATTATAAGGTCCTATT 817

RESULT 10
US-10-085-117-154/c
; Sequence 154, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154
; LENGTH: 78072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-154

Query Match 56.1%; Score 23; DB 15; Length 78072;
Best Local Similarity 74.4%; Pred. No. 2e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CTCTCCTGACTTTGTCGTAATAATTGATAAGATCTAAATT 40
||||| | ||||| ||||| ||||| ||||| |||||
Db 42822 CTCTCCTGACTCTAGTTGTAATGGAGCAGATGTAAC 42784

```

RESULT 11
US-10-282-122A-39149
; Sequence 39149, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39149
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-10-282-122A-39149

Query Match      54.6%; Score 22.4; DB 16; Length 2709;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      2      CTCTCCTGACTTTGTCCGTAAATTGATAAGATCTAATTG 41
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1309  CTCTCCGGACGCTGGCAGAAATCGATGAGATCCCGTTG 1348

RESULT 12
US-10-369-493-46473
; Sequence 46473, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

```

Best Local Similarity 78.8%; Pred. No. 1.6e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 CTGACTTTGTCCGTAATTTGATAAGATCTAATT 39
Db 186 CTGCCGCTGTCAATAAATTATAAGATTTAATT 154

RESULT 15
US-10-425-114-12000/c
; Sequence 12000, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12000
; LENGTH: 2047
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701152196_FLI
US-10-425-114-12000

Query Match 53.2%; Score 21.8; DB 16; Length 2047;
Best Local Similarity 78.8%; Pred. No. 2.3e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 9 GACTTTGTCCGTAATTTGATAAGATCTAATTG 41
Db 1172 GACTTTGACGGGAAATGATATGATTAGTTTG 1140

Search completed: November 3, 2004, 10:01:03
Job time : 77.6435 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 05:58:29 ; Search time 18.8325 Seconds
(without alignments)
1547.447 Million cell updates/sec

Title: US-09-933-267A-1_COPY_64311_64351

Perfect score: 41

Sequence: 1 tctctcctgacttcttcgt.....aattgataagatctaatttg 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.4	96.1	392000	4	US-10-027-983-11
2	23.8	58.0	1347	4	US-09-134-000C-389
3	21.4	52.2	497	4	US-09-270-767-6174
4	21.4	52.2	497	4	US-09-270-767-21456
5	21.2	51.7	395	2	US-08-967-101-19
6	21.2	51.7	395	2	US-08-592-541-19
7	21.2	51.7	395	3	US-09-124-698-19
8	21.2	51.7	395	3	US-09-127-480-19
9	21.2	51.7	395	3	US-08-496-841C-19
10	21.2	51.7	395	3	US-09-124-523-19
11	21.2	51.7	395	4	US-09-636-796A-19
12	21.2	51.7	395	4	US-08-431-048F-19
13	21	51.2	465	4	US-09-248-796A-5348
14	20.8	50.7	237	4	US-09-248-796A-1647
15	20.8	50.7	951	4	US-09-543-681A-1701
16	20.8	50.7	1557	3	US-07-956-483-25
17	20.8	50.7	1557	3	US-08-472-240A-17
18	20.8	50.7	2238	4	US-09-792-024-34
19	20.8	50.7	8339	4	US-08-956-171E-175
20	20.8	50.7	8339	4	US-08-781-986A-175
21	20.6	50.2	379	4	US-09-513-999C-15941
22	20.6	50.2	8549	4	US-08-956-171E-5
23	20.6	50.2	8549	4	US-08-781-986A-5
24	20.4	49.8	389	4	US-09-702-705-1766
25	20.4	49.8	389	4	US-09-736-457-1766
26	20.4	49.8	389	4	US-09-671-325-1766
27	20.4	49.8	389	4	US-09-658-824-1766

C 28	20.4	49.8	1590	3	US-09-334-938-8	Sequence 8, Appli
C 29	20.4	49.8	2365	4	US-08-956-171E-95	Sequence 95, Appl
C 30	20.4	49.8	2365	4	US-08-781-986A-95	Sequence 95, Appl
C 31	20.2	49.3	494	4	US-09-621-976-1823	Sequence 1823, Ap
C 32	20.2	49.3	1032	4	US-09-543-681A-2287	Sequence 2287, Ap
C 33	20.2	49.3	2904	3	US-09-221-294-3	Sequence 3, Appli
C 34	20.2	49.3	3428	4	US-09-919-039-77	Sequence 77, Appl
C 35	20	48.8	201	4	US-09-248-796A-8344	Sequence 8344, Ap
C 36	20	48.8	522	4	US-09-248-796A-6933	Sequence 6933, Ap
C 37	20	48.8	603	4	US-09-248-796A-13707	Sequence 13707, A
C 38	20	48.8	849	4	US-09-248-796A-435	Sequence 435, App
C 39	19.8	48.3	265	3	US-08-905-223-102	Sequence 102, App
C 40	19.8	48.3	797	3	US-09-071-353-1	Sequence 1, Appli
C 41	19.8	48.3	797	4	US-09-426-326-1	Sequence 1, Appli
C 42	19.8	48.3	1320	4	US-09-248-796A-5717	Sequence 5717, Ap
C 43	19.6	47.8	588	4	US-09-220-132-101	Sequence 101, App
C 44	19.4	47.3	204	4	US-09-583-110-2281	Sequence 2281, Ap
C 45	19.4	47.3	204	4	US-09-583-110-2284	Sequence 2284, Ap

ALIGNMENTS

RESULT 1
US-10-027-983-11
; Sequence 11, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: ~~US-10-027-983~~
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 392000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 137740
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 137742
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (138122) ... (138221)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 145507
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 151967
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (151967) ... (1542066)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: unsure
; LOCATION: 154217
; OTHER INFORMATION: unknown
; NAME/KEY: misc feature
; LOCATION: (164037) ... (164136)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (174657) ... (174756)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (186224) ... (186323)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: (195242) ... (195341)
; OTHER INFORMATION: n = A,T,C or G

Search 5

no 70057

#1

NAME/KEY: unsure
LOCATION: 202703
OTHER INFORMATION: unknown
NAME/KEY: misc feature
LOCATION: (202771)...(202870)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (206246)...(215602)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (218126)...(218225)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (220360)...(220459)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (222717)...(222816)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (223981)...(224080)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (227487)...(227586)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (230157)...(230256)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (232299)...(232398)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (236552)...(236651)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (238789)...(248788)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: exon
LOCATION: (118288)...(119101)
OTHER INFORMATION: exon 1C
NAME/KEY: exon:intron junction
LOCATION: (151129)...(151130)
OTHER INFORMATION: exon 5:intron 5
NAME/KEY: exon:intron junction
LOCATION: (299248)...(299249)
OTHER INFORMATION: exon 9:intron 9
NAME/KEY: exon:intron junction
LOCATION: (348578)...(348579)
OTHER INFORMATION: exon 10:intron 10
NAME/KEY: intron
LOCATION: (348579)...(381838)
OTHER INFORMATION: intron 10
NAME/KEY: intron:exon junction
LOCATION: (386185)...(386186)
OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Query Match 96.1%; Score 39.4; DB 4; Length 392000;
Best Local Similarity 97.6%; Pred. No. 1.8e-06;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTTGATAAGATCTAAATTG 41
|||||
Db 12773 TCTCTCCTGACTTTGTCCGTAATTTGATAAGATCTAAATTG 12813
|||||

RESULT 2
US-09-134-000C-389/c
Sequence 389, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 389
LENGTH: 1347
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-389

Query Match 58.0%; Score 23.8; DB 4; Length 1347;
Best Local Similarity 80.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TCTCCTGACTTTGTCCGTAATTTGATAAGATCTAA 37
|||||
Db 1151 TCTCCTTTCATTGTCCGTAATTTGATAATTCCTAA 1117
|||||

RESULT 3
US-09-270-767-6174
Sequence 6174, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6174
LENGTH: 497
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-6174

Query Match 52.2%; Score 21.4; DB 4; Length 497;
Best Local Similarity 71.8%; Pred. No. 10;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTTGATAAGATCTAAATT 39
|||||
Db 330 TCTGCTCTGATTTCGTCCCAAACTGCAAAGATATATT 368
|||||

RESULT 4
US-09-270-767-21456
Sequence 21456, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21456
LENGTH: 497
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-21456

Query Match 52.2%; Score 21.4; DB 4; Length 497;
Best Local Similarity 71.8%; Pred. No. 10;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTTGATAAGATCTAAATT 39
|||||

Db 330 TCTGCTCTGATTTCGTCGCCCAAACTGCCAAAGATATAATT 368

RESULT 5

US-08-967-101-19
; Sequence 19, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 395 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-967-101-19

Query Match 51.7%; Score 21.2; DB 2; Length 395;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TCCTGACTTTGTCGTAATAATTGATAAGATCTAAT 38

Db 271 TTCTGAGTTTCTCCATAAAATTAATTGGACCTAAT 304

RESULT 6

US-08-592-541-19
; Sequence 19, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.

; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-592-541-19

Query Match 51.7%; Score 21.2; DB 2; Length 395;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TCCTGACTTTGTCGTAATAATTGATAAGATCTAAT 38

Db 271 TTCTGAGTTTCTCCATAAAATTAATTGGACCTAAT 304

RESULT 7

US-09-124-698-19
; Sequence 19, Application US/09124698
; Patent No. 6117978
; GENERAL INFORMATION:

APPLICANT: ST. GEORGE-HYSLOP, PETER H

APPLICANT: ROMMENS, JOHANNA M

APPLICANT: FRASER, PAUL E

TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED

TITLE OF INVENTION: TO ALZHEIMER'S DISEASE

NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: High Street Tower - 125 High Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/124,698

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/592,541

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 395 base pairs

TYPE: nucleic acid

STRANDEDNESS: single


```

; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-124-523-19
Query Match 51.7%; Score 21.2; DB 3; Length 395;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TCCTGACTTTGTCGTAATGATAAGATCTAAT 38
Db 271 TTCTGAGTTTCTCCATAAATAATTGGACCTAAT 304

RESULT 11
US-09-636-796A-19
; Sequence 19, Application US/09636796A
; Patent No. 6485911
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/636,796A
; FILING DATE: 11-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-636-796A-19
Query Match 51.7%; Score 21.2; DB 4; Length 395;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 5 TCCTGACTTTGTCGTAATGATAAGATCTAAT 38
Db 271 TTCTGAGTTTCTCCATAAATAATTGGACCTAAT 304

RESULT 12
US-08-431-048F-19
; Sequence 19, Application US/08431048F
; Patent No. 6531586
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 155
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY P.C.
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,048F
; FILING DATE: 28-Apr-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: FEHLNER, PAUL F.
; REGISTRATION NUMBER: 35135
; REFERENCE/DOCKET NUMBER: 1034/0F808
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-527-6237
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-431-048F-19
Query Match 51.7%; Score 21.2; DB 4; Length 395;
Best Local Similarity 76.5%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TCCTGACTTTGTCGTAATGATAAGATCTAAT 38
Db 271 TTCTGAGTTTCTCCATAAATAATTGGACCTAAT 304

RESULT 13
US-09-248-796A-5348
; Sequence 5348, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409

```

; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5348
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5348

Query Match 51.2%; Score 21; DB 4; Length 465;
Best Local Similarity 73.0%; Pred. No. 14;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 TCTGACTTTGTCCGTAATTTGATAAGATCTAATTTG 41
Db 306 TACTGAGATGGTCACTCAATTCATTCATCTAATTTG 342

RESULT 14
US-09-248-796A-1647/c
; Sequence 1647, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1647
; LENGTH: 237
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1647

Query Match 50.7%; Score 20.8; DB 4; Length 237;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTCGCGTAAATTTGATAAGATCTAATTT 40
Db 102 TTTATCTTTATTTTCGTTCAATTCCTAAATATAATTT 63

RESULT 15
US-09-543-681A-1701/c
; Sequence 1701, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1701
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1701

Query Match 50.7%; Score 20.8; DB 4; Length 951;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTCGCGTAAATTTGATAAGATCTAATTT 40

Db 46 TCACTCCTGATGTTTCAGTTAAATTTGATATTAGCTAATT 7

Search completed: November 3, 2004, 09:46:50
Job time : 19.8325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:55:34 ; Search time 255.809 Seconds
(without alignments)
7579.407 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306362_306402

Perfect score: 41

Sequence: 1 ttgttgatgctgagcccc.....atatctattccgagtatga 41

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	392	9 AF120105	AF120105 Homo sapi
2	41	100.0	1218	6 AR397495	AR397495 Sequence
3	41	100.0	1218	9 AF172068	AF172068 Homo sapi
4	41	100.0	1218	9 AF172069	AF172069 Homo sapi
5	41	100.0	1223	6 AX066401	AX066401 Sequence
6	41	100.0	1237	9 HSESR13	AF123496 Homo sapi
7	41	100.0	1374	6 AX066402	AX066402 Sequence
8	41	100.0	1644	12 AF242866	AF242866 Synthetic
9	41	100.0	1770	6 AX824420	AX824420 Sequence
10	41	100.0	1788	6 AX411751	AX411751 Sequence
11	41	100.0	1803	6 AX824417	AX824417 Sequence
12	41	100.0	1956	6 AR116270	AR116270 Sequence
13	41	100.0	1983	6 AX411753	AX411753 Sequence
14	41	100.0	1983	6 AX411755	AX411755 Sequence
15	41	100.0	1983	6 AX411757	AX411757 Sequence
16	41	100.0	2092	6 I08538	I08538 Sequence 1
17	41	100.0	2092	6 AX474709	AX474709 Sequence
18	41	100.0	2092	9 HUMERMCF	M12674 Humanestro
19	41	100.0	2106	9 HSU47678	U47678 Human 80 kD

20	41	100.0	2220	6 BD016996	BD016996 Fused est
21	41	100.0	2322	6 A42099	A42099 Sequence 1
22	41	100.0	2532	6 BD187708	BD187708 Specific
23	41	100.0	4963	6 AR029418	AR029418 Sequence
24	41	100.0	4963	6 I15368	I15368 Sequence 16
25	41	100.0	4963	6 I56760	I56760 Sequence 9
26	41	100.0	5439	9 HSM807087	BX640939 Homo sapi
27	41	100.0	6039	6 CQ832274	CQ832274 Sequence
28	41	100.0	6450	6 AR153585	AR153585 Sequence
29	41	100.0	6450	6 AR225684	AR225684 Sequence
30	41	100.0	6450	6 AR397400	AR397400 Sequence
31	41	100.0	6450	6 AX587952	AX587952 Sequence
32	41	100.0	6450	9 HSERR	X03635 Homo sapien
33	41	100.0	6567	6 AX128350	AX128350 Sequence
34	41	100.0	6610	12 AF061181	AF061181 Mammalian
35	41	100.0	6623	6 AX128345	AX128345 Sequence
36	41	100.0	6639	6 AX128351	AX128351 Sequence
37	41	100.0	6695	6 AX128347	AX128347 Sequence
38	41	100.0	6695	6 AX128353	AX128353 Sequence
39	41	100.0	6695	6 AX128354	AX128354 Sequence
40	41	100.0	6746	6 AX128344	AX128344 Sequence
41	41	100.0	6801	6 AX128352	AX128352 Sequence
42	41	100.0	6801	6 AX128355	AX128355 Sequence
43	41	100.0	6818	6 AX128346	AX128346 Sequence
44	41	100.0	6828	6 AX128340	AX128340 Sequence
45	41	100.0	6833	6 AX128349	AX128349 Sequence

ALIGNMENTS

RESULT 1
AF120105
LOCUS AF120105 392 bp mRNA linear PRI 06-APR-1999
DEFINITION Homo sapiens alternatively-spliced estrogen receptor alpha mRNA,
partial cds.
ACCESSION AF120105
VERSION AF120105.1 GI:4567039
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Campbell-Thompson,M.L.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Medicine, University of Florida, Box 100214
Gastroenterology, Gainesville, FL 32667, USA

FEATURES

source
1..392
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q25.1"
/sex="male"
/cell_type="colon adenocarcinoma"
CDS
<1..343
/note="steroid hormone receptor; missing exon 5"
/codon_start=2
/product="alternatively-spliced estrogen receptor alpha"
/protein_id="AAD23565.1"
/db_xref="GI:4567040"
/translation="RRGGRMLKHQRDDGEGRGEVGSAGDMRAANLWPSPLMKRSK
KNSLALSLTADQMVSALLDAEPIPLYSEYDTPRFSEASMMGLLTNLADELVHMV
AKRVPGTRENV"
variation
25
/note="compared to wild type sequence deposited in GenBank
Accession Number X03635"
/replace="a"

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 392;

Best Local Similarity 100.0%; Pred. No. 8.5e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 182 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 222

RESULT 2
AR397495
LOCUS AR397495 1218 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 98 from patent US 6617162.
ACCESSION AR397495
VERSION AR397495.1 GI:40134376
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1218)
AUTHORS Dobie,K.W. and Roach,M.P.
TITLE Antisense modulation of estrogen receptor alpha expression
JOURNAL Patent: US 6617162-A 98 09-SEP-2003;
FEATURES Location/Qualifiers
source
1. .1218
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1218;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425

RESULT 3

AF172068
LOCUS AF172068 1218 bp mRNA linear PRI 02-OCT-2000
DEFINITION Homo sapiens fatty acid synthase/estrogen receptor fusion protein
(FAS/ER) mRNA, partial cds.

ACCESSION AF172068
VERSION AF172068.1 GI:5825505
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1218)
AUTHORS Ye,Q., Chung,L.W., Li,S. and Zhau,H.E.
TITLE Identification of a novel FAS/ER-alpha fusion transcript expressed in human cancer cells

JOURNAL Biochim. Biophys. Acta 1493 (3), 373-377 (2000)

MEDLINE 20472057
PUBMED 11018265

REFERENCE 2 (bases 1 to 1218)
AUTHORS Ye,Q.N., Li,S.W., Chung,L.W.K. and Zhau,H.Y.

TITLE Direct Submission
JOURNAL Submitted (25-JUL-1999) Department of Urology, University of Virginia, 1300 Jefferson Park Avenue, Charlottesville, VA 22908, USA

FEATURES
source

Location/Qualifiers
1. .1218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="393"
/cell_line="ARCap"
/cell_type="prostate cancer"
<1. .1218
/gene="FAS/ER"
1. .1218

gene

CDS

/gene="FAS/ER"
/codon_start=1
/product="fatty acid synthase/estrogen receptor fusion protein"
/protein_id="AAD53275.1"
/db_xref="GI:5825506"
/translation="SLARTSSAEYNNVNLVSPVLFQEALWHVPEHAVVLEIAPHALLQ
AVLKRGLKPSCTIIPLMKKDRDNLEFFLAGIRRLHLSGIDANPNALFPVVEFPAPRG
TPLISPLIKWDHSLALSLTADQMVSAALLDAEPPILYSEYDPTFPFSEASMMGLLTNLA
DRELVHMINKAKRVPGFVDLTLDHQVHLECAWLEILMIGLVWRSMHPGKLLFAPNL
LLDRNQKCVGEMVEIFDMLLATSSRFMMNLQGEFVCLKSIILLNSGVYTFLSSTL
KSLEEKDHIHRVLDKITDTLIHLMAKAGLTQQQHQRLAQLLLILSHIRHMSNKGMEH
LYSMKCKNVVPLYDLLLEMLDAHRLHAPTSRGGASVEETDQSHLATAGSTSSHSLOKY
YITGEAEGFPATV"

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 1218;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
|||||
Db 385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425

RESULT 4

AF172069
LOCUS AF172069 1218 bp mRNA linear PRI 06-SEP-1999
DEFINITION Homo sapiens fatty acid synthase/estrogen receptor fusion protein
(FAS/ER) mRNA, partial cds.

ACCESSION AF172069
VERSION AF172069.1 GI:5825508
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1218)
AUTHORS Ye,Q.N., Li,S.W., Chung,L.W.K. and Zhau,H.Y.

TITLE Direct Submission
JOURNAL Submitted (25-JUL-1999) Department of Urology, University of Virginia, 1300 Jefferson Park Avenue, Charlottesville, VA 22908, USA

FEATURES
source

Location/Qualifiers
1. .1218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="73"
/cell_line="LNCap"
/cell_type="prostate cancer"
<1. .1218
/gene="FAS/ER"
1. .1218
/gene="FAS/ER"
/codon_start=1
/product="fatty acid synthase/estrogen receptor fusion protein"
/protein_id="AAD53276.1"
/db_xref="GI:5825509"

gene

CDS

FEATURES
source

Location/Qualifiers
1. .1218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="73"
/cell_line="LNCap"
/cell_type="prostate cancer"
<1. .1218
/gene="FAS/ER"
1. .1218
/gene="FAS/ER"
/codon_start=1
/product="fatty acid synthase/estrogen receptor fusion protein"
/protein_id="AAD53276.1"
/db_xref="GI:5825509"

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 1218;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 385 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 425

RESULT 5
AX066401
LOCUS 1223 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 1 from Patent WO0100823.
ACCESSION AX066401
VERSION AX066401.1 GI:12544111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Gannon,F., Denger,S. and Flouriot,G.
TITLE Novel isoforms of the human estrogen receptor_g(a)
JOURNAL Patent: WO 0100823-A 1 04-JAN-2001;
EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
FEATURES
source Location/Qualifiers
1. .1223
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 41; DB 6; Length 1223;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 390 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 430

RESULT 6
HSESRI3
LOCUS 1237 bp DNA linear PRI 07-APR-2000
DEFINITION Homo sapiens estrogen receptor alpha (ESR1) gene, exon 4.
ACCESSION AF123496
VERSION AF123496.1 GI:5821720
KEYWORDS
SEGMENT
SOURCE 3 of 7
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Schubert,E.L., Lee,M.K., Newman,B. and King,M.C.
MEDLINE Single nucleotide polymorphisms (SNPs) in the estrogen receptor
PUBMED gene and breast cancer susceptibility
20084372 J. Steroid Biochem. Mol. Biol. 71 (1-2), 21-27 (1999)
10619354
REFERENCE 2 (bases 1 to 1237)
AUTHORS Schubert,E.L., Lee,M.K. and King,M.-C.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1999) Medical Genetics, University of Washington,
1959 NE Pacific Street, Seattle, WA 98195-7720, USA
FEATURES
source Location/Qualifiers
1. .1237
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q25"
350. .685
/gene="ESR1"
/number=4

exon

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 1237;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 544 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 584

RESULT 7
AX066402
LOCUS 1374 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 2 from Patent WO0100823.
ACCESSION AX066402
VERSION AX066402.1 GI:12544112
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Gannon,F., Denger,S. and Flouriot,G.
FEATURES Novel isoforms of the human estrogen receptor_g(a)
source Patent: WO 0100823-A 2 04-JAN-2001;
1. .1374 EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 41; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 541 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 581

RESULT 8
AF242866
LOCUS 1644 bp mRNA linear SYN 24-APR-2000
DEFINITION Synthetic construct ECFP-estrogen receptor ligand binding domain
ACCESSION AF242866
VERSION AF242866.1 GI:7638248
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE artificial sequences.
1 (bases 1 to 1644)
AUTHORS Llopis,J., Westin,S., Ricote,M., Wang,J., Cho,C.Y., Kurokawa,R.,
TITLE Mullen,T.M., Rose,D.W., Rosenfeld,M.G., Tsien,R.Y. and Glass,C.K.
Ligand-dependent interactions of coactivators steroid receptor
coactivator-1 and peroxisome proliferator-activated receptor
binding protein with nuclear hormone receptors can be imaged in
live cells and are required for transcription
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (B), 4363-4368 (2000)
MEDLINE 20226109
PUBMED 10760302
REFERENCE 2 (bases 1 to 1644)
AUTHORS Llopis,J., Westin,S., Ricote,M., Wang,J., Cho,C.Y., Kurokawa,R.,
TITLE Mullen,T.M., Rose,D.W., Rosenfeld,M.G., Tsien,R.Y. and Glass,C.K.
JOURNAL Direct Submission
FEATURES Submitted (09-MAR-2000) Medicine/Cellular and Molecular Medicine,
source UC San Diego, 9500 Gilman Drive, La Jolla, CA 92103, USA
1. .1644
/organism="synthetic construct"
/mol_type="mRNA"

source /db_xref="taxon:32630"
/focus
697. .1644
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .1644
/codon_start=1
/transl_table=11
/product="ECFP-estrogen receptor ligand binding domain
fusion protein"
/protein_id="AAF65451.1"
/db_xref="GI:7638249"
/translation="MVSKGEELFTGVVPILVELDGDVNGHRFSVSGEGDATYGKLT
LKFICTTGKLPVPWPTLTITLTGWQCFSRYPDHMKQDFFKSAPEGVYQERTIFFK
DDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNIISHNVIYITADKQKN
GIKAHFKIRHNIEDGSVQLADHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDRH
MVLLEFVTAARNPFGSGAGDMRAANLWPSPLMIKRSKNSLALSLTADQMVSAALLDAEP
PILYSEYDPTRPFSSEASMMGLLTNLADRELVMINWAKRVPGFVDTLHDQVHLLLECA
WLEILMIGLVWRSMHPVKLLFAPNLLDRNQKCEGMVEIFDMLLATSSRFMRMNL
QGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLO
QQHORLAQLLLSHIRHMSNKGMEHLYSMKCKNVVPLYDILLLEMLDAHRLHAPTSGR
GASVEETDQSHLATAGTSSSHLSLQKYYITGEAEGFPATV"

ORIGIN

Query Match 100.0%; Score 41; DB 12; Length 1644;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 811 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 851

RESULT 9
AX824420
LOCUS AX824420 1770 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 6 from Patent WO03070975.
ACCESSION AX824420
VERSION AX824420.1 GI:39750420
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Verdonk,G.N., Dijkema,R.N. and Schoonen,W.G.
TITLE An estrogen receptor alpha construct for use in a yeast 2-hybrid assay
JOURNAL Patent: WO 03070975-A 6 28-AUG-2003;
Akzo Nobel N.V. (NL)
FEATURES Location/Qualifiers
source 1. .1770
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion DNA sequence page 17-19"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1770;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 937 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 977

RESULT 10
AX411751
LOCUS AX411751 1788 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0228175.
ACCESSION AX411751
VERSION AX411751.1 GI:21444272

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chambon,P. and Metzger,D.
TITLE Transgenic mouse for targeted recombination mediated by modified Cre-er
JOURNAL Patent: WO 0228175-A 1 11-APR-2002;
ASS POUR LE DEV DE LA RECH (FR)
FEATURES Location/Qualifiers
source 1. .1788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .1788
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35258.1"
/db_xref="GI:21444273"
/translation="MTMTLHTKASGMALLHQIQNELEPLNRPQLKIPLERPLGEVYL
DSSKPAVNYPEGAAYEFNAAAANAQVYQGTGLPYGPGSEAAAFSGNGLGFPPLNS
VSPPLMLLHPPQLSPFLQPHGQQVYFYLENEPSTGYTVREAGPPAFYRPSNDNRQOG
GRERLASTNDKGSMAESAKETRYCAVNDYASGYHYGWSCEGCKAFFKRSIQGHND
YMCPATNQCTIDKNRRKSCQACRLKCEYGVMMKGIRKDRRGGRMLKHKFQDDGEG
RGEVGSAGDMRAANLWPSPLMIKRSKNSLALSLTADQMVSAALLDAEPPILYSEYDPT
RPFSEASMMGLLTNLADRELVMINWAKRVPGFVDTLHDQVHLLLECAWLEIIMIGLV
WRSMEHPGKLLFAPNLLDRNQKCEGMVEIFDMLLATSSRFMRMNLQGEFVCLKS
IILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLOQQHORLAQLL
LILSHIRHMSNKGMEHLYSMKCKNVVPLYDILLLEMLDAHRLHAPTSGRGGASVEETDQS
HLATAGTSSSHLSLQKYYITGEAEGFPATV"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1788;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 955 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 995

RESULT 11
AX824417
LOCUS AX824417 1803 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 3 from Patent WO03070975.
ACCESSION AX824417
VERSION AX824417.1 GI:39750419
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Verdonk,G.N., Dijkema,R.N. and Schoonen,W.G.
TITLE An estrogen receptor alpha construct for use in a yeast 2-hybrid assay
JOURNAL Patent: WO 03070975-A 3 28-AUG-2003;
Akzo Nobel N.V. (NL)
FEATURES Location/Qualifiers
source 1. .1803
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion DNA p 14-16"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1803;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41

Db 970 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1010

RESULT 12
AR116270 LOCUS AR116270 1956 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6133027.
ACCESSION AR116270
VERSION AR116270.1 GI:14096592
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1956)
AUTHORS Yee,J.-K., Friedmann,T. and Chen,S.-T.
TITLE Inducible expression system
JOURNAL Patent: US 6133027-A 17-OCT-2000
FEATURES
source
1. .1956
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 1123 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1163

RESULT 13
AX411753 LOCUS AX411753 1983 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3 from Patent WO0228175.
ACCESSION AX411753
VERSION AX411753.1 GI:21444274
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Chambon,P. and Metzger,D.
TITLE Transgenic mouse for targeted recombination mediated by modified Cre-er
JOURNAL Patent: WO 0228175-A 3 11-APR-2002;
ASS POUR LE DEV DE LA RECH (FR)
FEATURES
source
1. .1983
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1. .1983
/note="unnamed protein product; Chimeric sequence
Homosapiens-Bacteriophage P1"
/codon_start=1
/transl_table=11
/protein_id="CAD35259.1"
/db_xref="GI:21444275"
/translation="MSNLLTVHQNLPAIPVDTSDVVRKNLMDMFRDQAFSEHTWKMLLSVCRSAAWCKLNNRKWFPAPEDVRDYLILQARGLAVKTIQQHLGQLNMLHRRSGLPRPDSNAVSLVMRRIRKENVDAGERAKQALAFERTDFQVRSMLMENSDDRCQDIRNLAFLGIAYNTLLRIAEIARIVKDISRTDGGRLIHIGRTKTLVSTAGVEKALSIGVTKLVERWISVGVAADPNNYLFCRVKNGVAAPSATSQLSSTRALEGIFEATHRLIYGAKDDSGORYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTNNIVMNYIRNLDSETGAMVRLLLEDGDLPSAGDMAANLWPSPLMIKRSKNSLALSALTADQMVSAILDAPPPILYSEYDPTRPFSSEASMMGLLTNADRELVHMINWAKRVPGFVDLTLHDQVHLLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDNRNQKCVGEMVEIFDMLLATSSRRFRMNLQGE EFVCLKSIILLNSGVYTFLSSTLKSLEEKDHHRVLDKITDTLIHLMKAGLTQQQH QRLAQLLLILSHIRHMSNKRMEHLYSMCKKNVVPPLYDILLLEMLDAHRLHAPTSRGGAS VEETDQSHLATAGTSSSHSLQKYYITGEAEGFPATV"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 1123 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1163

RESULT 14
AX411755 LOCUS AX411755 1983 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 5 from Patent WO0228175.
ACCESSION AX411755
VERSION AX411755.1 GI:21444276
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Chambon,P. and Metzger,D.
TITLE Transgenic mouse for targeted recombination mediated by modified Cre-er
JOURNAL Patent: WO 0228175-A 5 11-APR-2002;
ASS POUR LE DEV DE LA RECH (FR)
FEATURES
source
1. .1983
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1. .1983
/note="unnamed protein product; Chimeric sequence
Homosapiens-Bacteriophage P1"
/codon_start=1
/transl_table=11
/protein_id="CAD35260.1"
/db_xref="GI:21444277"
/translation="MSNLLTVHQNLPAIPVDTSDVVRKNLMDMFRDQAFSEHTWKMLLSVCRSAAWCKLNNRKWFPAPEDVRDYLILQARGLAVKTIQQHLGQLNMLHRRSGLPRPDSNAVSLVMRRIRKENVDAGERAKQALAFERTDFQVRSMLMENSDDRCQDIRNLAFLGIAYNTLLRIAEIARIVKDISRTDGGRLIHIGRTKTLVSTAGVEKALSIGVTKLVERWISVGVAADPNNYLFCRVKNGVAAPSATSQLSSTRALEGIFEATHRLIYGAKDDSGORYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTNNIVMNYIRNLDSETGAMVRLLLEDGDLPSAGDMAANLWPSPLMIKRSKNSLALSALTADQMVSAILDAPPPILYSEYDPTRPFSSEASMMGLLTNADRELVHMINWAKRVPGFVDLTLHDQVHLLLECAWLEILMIGLVWRSMEHPVKLLFAPNLLLDNRNQKCVGEMVEIFDMLLATSSRRFRMNLQGE EFVCLKSIILLNSGVYTFLSSTLKSLEEKDHHRVLDKITDTLIHLMKAGLTQQQH QRLAQLLLILSHIRHMSNKRMEHLYSMCKKNVVPPLYDILLLEMLDAHRLHAPTSRGGAS VEETDQSHLATAGTSSSHSLQKYYITGEAEGFPATA"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 1983;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 41
Db 1150 TTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGA 1190

RESULT 15
AX411757 LOCUS AX411757 1983 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 7 from Patent WO0228175.
ACCESSION AX411757
VERSION AX411757.1 GI:21444278
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Chambon,P. and Metzger,D.

ORIGIN

Search completed: November 3, 2004, 07:43:07
Job time : 257.809 secs

Search

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:55:34 ; Search time 255.809 Seconds
(without alignments)
7579.407 Million cell updates/sec

Title: US-09-933-267A-1_COPY_64311_64351

Perfect score: 41

Sequence: 1 tctctcgtgactttgtccgt.....aattgataagatctaatttg 41

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	100.0	56124	9	AB090237	AB090237 Homo sapi
2	41	100.0	180832	2	AC044828	AC044828 Homo sapi
3	41	100.0	349980	6	AX232503	AX232503 Sequence
4	41	100.0	349980	6	AX453703	AX453703 Sequence
5	39.4	96.1	110000	6	AR397408_0	AR397408 Sequence
6	39.4	96.1	119737	9	HS404G5	AL035695 Human DNA
7	26.6	64.9	210954	5	AJ627213	AJ627213 Gallus ga
8	25	61.0	179906	2	AC119339	AC119339 Rattus no
9	24	58.5	614	14	AF065554	AF065554 HIV-1 iso
10	24	58.5	131823	2	AC010497	AC010497 Homo sapi
11	24	58.5	134349	9	AC010296	AC010296 Homo sapi
12	24	58.5	154061	9	AC008524	AC008524 Homo sapi
13	24	58.5	179647	2	AC108124	AC108124 Homo sapi
14	24	58.5	188207	2	AC010233	AC010233 Homo sapi
15	23.8	58.0	1347	6	AR394374	AR394374 Sequence
16	23.8	58.0	86761	10	AL663044	AL663044 Mouse DNA
17	23.8	58.0	213038	2	AC111823	AC111823 Rattus no
18	23.8	58.0	235502	2	AC103222	AC103222 Rattus no
19	23.8	58.0	300225	1	AE016949	AE016949 Enterococ

C 20	23.6	57.6	151341	10	AC115236	AC115236 Rattus no
C 21	23.6	57.6	181157	10	AC121789	AC121789 Mus muscu
22	23.6	57.6	207528	2	AC139331	AC139331 Mus muscu
23	23.6	57.6	245877	2	AC128382	AC128382 Rattus no
24	23.6	57.6	260121	2	AC130738	AC130738 Rattus no
25	23.6	57.6	270188	2	AC112821	AC112821 Rattus no
26	23.6	57.6	338614	2	AC103268	AC103268 Rattus no
27	23.4	57.1	591	6	CQ103678	CQ103678 Sequence
28	23.4	57.1	591	6	CQ142508	CQ142508 Sequence
29	23.4	57.1	591	6	CQ225743	CQ225743 Sequence
30	23.4	57.1	591	6	CQ263783	CQ263783 Sequence
31	23.4	57.1	591	6	CQ300858	CQ300858 Sequence
32	23.4	57.1	10585	1	AE006124	AE006124 Pasteurel
33	23.4	57.1	110000	2	BX294176_3	Continuation (4 of
34	23.4	57.1	150831	9	AC007719	AC007719 Homo sapi
35	23.4	57.1	171263	10	AC118702	AC118702 Mus muscu
36	23.4	57.1	184674	2	AC102557	AC102557 Mus muscu
37	23.4	57.1	192449	10	AC115884	AC115884 Mus muscu
38	23.4	57.1	207870	10	AL669900	AL669900 Mouse DNA
39	23.4	57.1	235317	2	AC121371	AC121371 Rattus no
40	23.4	57.1	238007	2	AC107569	AC107569 Rattus no
41	23.4	57.1	241571	2	AC108281	AC108281 Rattus no
42	23.2	56.6	123947	9	AC005045	AC005045 Homo sapi
43	23.2	56.6	168745	2	AC119552	AC119552 Rattus no
44	23.2	56.6	248329	2	AC109989	AC109989 Rattus no
45	23	56.1	609	11	BV079087	BV079087 px-61f6 F

ALIGNMENTS

RESULT 1
AB090237
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB090237
Homo sapiens gene for estrogen receptor alpha, exon 1a, 1b, 2.
AB090237
AB090237.1 GI:28460664

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

Inoue,A., Hayashi,S., Aoyagi,K., Nishigaki,M., Sasaki,H. and Kiyama,R.

A reporter gene assay for evaluation of tissue-specific responses to estrogens based on the differential use of promoters A to F of the human estrogen receptor alpha gene

Unpublished

22516287

12628303

2 (bases 1 to 56124)

Sasaki,H.

Direct Submission

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted 20-AUG-2002

Research Institute, Genetics Division; 1-1, Tsukiji 5-chome,

Chuo-ku, Tokyo 104-0045, Japan (E-mail:hksasaki@gan2.res.ncc.go.jp,

Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)

Location/Qualifiers

1. .56124

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

join(7128..7219,53065..>53195)

/product="estrogen receptor alpha"

/note="generated by promoter E and alternative splicing"

7128..7219

/product="estrogen receptor alpha"

join(41744..41866,53065..>53195)

/product="estrogen receptor alpha"

/note="generated by promoter F and alternative splicing"

41744..41866

/product="estrogen receptor alpha"
/number=1b
53065..53195
/product="estrogen receptor alpha"
/number=2

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 56124;
Best Local Similarity 100.0%; Pred. No. 8.le-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTCTGACTTTGTCGTAATGATAAGATCTAATTG 41
|||||
Db 53226 TCTCTGACTTTGTCGTAATGATAAGATCTAATTG 53266
|||||

RESULT 2

AC044828/c
LOCUS 180832 bp DNA linear HTG 20-SEP-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-282P11 map 6, WORKING-DRAFT
SEQUENCE, 34 unordered pieces.

AC044828

AC044828.2 GI:10198537

HTG; HTGS PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 180832)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 6, clone RP11-282P11

Unpublished

2 (bases 1 to 180832)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,

Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,

Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,

Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,

Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,

McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 20, 2000 this sequence version replaced gi:7543791.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L8038

Center clone name: 282_P_11

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 162722 bases at least Q40
Consensus quality: 171198 bases at least Q30
Consensus quality: 174912 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 177532; sum-of-contigs
Quality coverage: 4.1 in Q20 bases; agarose-fp
Quality coverage: 4.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

3320: contig of 3320 bp in length
3420: gap of 100 bp
3421
3421
4535: contig of 1115 bp in length
4635: gap of 100 bp
4636
5881: contig of 1246 bp in length
5882
5882
5882
7889: contig of 1908 bp in length
7890
7890
7990
9338: contig of 1349 bp in length
9339
9438: gap of 100 bp
9439
10913: contig of 1475 bp in length
10914
11013: gap of 100 bp
11014
13206: contig of 2193 bp in length
13207
13306: gap of 100 bp
13307
14999: contig of 1693 bp in length
15000
15099: gap of 100 bp
18000: contig of 2901 bp in length
18001
18100: gap of 100 bp
18101
20868: contig of 2768 bp in length
20869
20968: gap of 100 bp
20969
24646: contig of 3678 bp in length
24647
24746: gap of 100 bp
24747
28762: contig of 4016 bp in length
28763
28862: gap of 100 bp
31776: contig of 2914 bp in length
31777
31876: gap of 100 bp
31877
35418: contig of 3542 bp in length
35419
35518: gap of 100 bp
40231: contig of 4713 bp in length
40331: gap of 100 bp
40332
44448: contig of 4117 bp in length
44449
44548: gap of 100 bp
49300: contig of 4752 bp in length
49301
49400: gap of 100 bp
53231: contig of 3831 bp in length
53331: gap of 100 bp
53332
57697: contig of 4366 bp in length
57698
57797: gap of 100 bp
62107: contig of 4310 bp in length
62108
62207: gap of 100 bp
62208
65070: contig of 2863 bp in length
65071
65170: gap of 100 bp
70534: contig of 5364 bp in length
70535
70634: gap of 100 bp
77999: contig of 7365 bp in length
78000
78099: gap of 100 bp
83750: contig of 5651 bp in length
83751
83850: gap of 100 bp
90537: contig of 6687 bp in length
90538
90638
97383: contig of 6746 bp in length
97384
97483: gap of 100 bp
106001: contig of 8518 bp in length
106101: gap of 100 bp
106102
116583: contig of 10482 bp in length
116584
116683: gap of 100 bp
125776: contig of 9033 bp in length

* 125777 125876: gap of 100 bp
* 125877 131540: contig of 5664 bp in length
* 131541 131640: gap of 100 bp
* 131641 141557: contig of 9917 bp in length
* 141558 141657: gap of 100 bp
* 141658 154976: contig of 13319 bp in length
* 154977 155076: gap of 100 bp
* 155077 167892: contig of 12816 bp in length
* 167893 167992: gap of 100 bp
* 167993 180832: contig of 12840 bp in length.

FEATURES
source
1. .180832
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6"
/clone="RP11-282P11"
/clone_lib="RPCI-11 Human Male BAC"
1. .3320
/note="assembly_fragment"
clone_end:SP6
vector_side:left
3421. .4535
/note="assembly_fragment"
4636. .5881
/note="assembly_fragment"
5982. .7889
/note="assembly_fragment"
7990. .9338
/note="assembly_fragment"
9439. .10913
/note="assembly_fragment"
11014. .13206
/note="assembly_fragment"
13307. .14999
/note="assembly_fragment"
15100. .18000
/note="assembly_fragment"
18101. .20868
/note="assembly_fragment"
20969. .24646
/note="assembly_fragment"
24747. .28762
/note="assembly_fragment"
28863. .31776
/note="assembly_fragment"
31877. .35418
/note="assembly_fragment"
35519. .40231
/note="assembly_fragment"
40332. .44448
/note="assembly_fragment"
44549. .49300
/note="assembly_fragment"
49401. .53231
/note="assembly_fragment"
53332. .57697
/note="assembly_fragment"
57798. .62107
/note="assembly_fragment"
62208. .65070
/note="assembly_fragment"
65171. .70534
/note="assembly_fragment"
70635. .77999
/note="assembly_fragment"
78100. .83750
/note="assembly_fragment"
83851. .90537
/note="assembly_fragment"
90638. .97383

Query Match 100.0%; Score 41; DB 2; Length 180832;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
|||||
Db 108986 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 108946
|||||

RESULT 3
AX232503
LOCUS AX232503 349980 bp DNA linear PAT 11-SEP-2001
DEFINITION Sequence 1 from Patent WO0162969.
ACCESSION AX232503
VERSION AX232503.1 GI:15592567
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kalush,F., Cassel,M.J., Hwang,S.S. and Winn-Deen,E.S.
TITLE Estrogen receptor alpha variants and methods of detection thereof
JOURNAL Patent: WO 0162969-A 1 30-AUG-2001;
PE Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1. .349980
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Original sequence 1: Length 465237 has been
replaced by~seq 1: from 1 to 349980 349980 bases~seq 3:
from 300001 to 465237 165237 bases"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 41
|||||
Db 64311 TCTCTCCTGACTTTGTCCGTAATAATTGATAAGATCTAATTG 64351
|||||

RESULT 4
AX453703
LOCUS AX453703 349980 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0234945.
ACCESSION AX453703
VERSION AX453703.1 GI:21712921
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kalush,F., Cassel,M.J., Hwang,S.S. and Winn-Deen,E.S.
TITLE Estrogen receptor alpha variants and methods of detection thereof
JOURNAL Patent: WO 0234945-A 1 02-MAY-2002;
Applera Corporation (US)
FEATURES
source
Location/Qualifiers
1. .349980
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="original length seq 1: 465237~replaced by seq 1:
000.001 -> 349.980~seq 4: 300.001 -> 465237"

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 41
|||||
Db 64311 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 64351
|||||

RESULT 5
AR397408_0
WPCOMMENT
Sequence split into 4 fragments LOCUS AR397408 Accession AR397408
Fragment Name Begin End
AR397408_0 1 110000
AR397408_1 100001 210000
AR397408_2 200001 310000
AR397408_3 300001 392000
LOCUS AR397408 392000 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 11 from patent US 6617162.
ACCESSION AR397408
VERSION AR397408.1 GI:40134206
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 392000)
AUTHORS Dobie,K.W. and Roach,M.P.
TITLE Antisense modulation of estrogen receptor alpha expression
JOURNAL Patent: US 6617162-A 11 09-SEP-2003;
FEATURES Location/Qualifiers
source 1. .392000
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 96.1%; Score 39.4; DB 6; Length 110000;
Best Local Similarity 97.6%; Pred. No. 0.00031;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 41
|||||
Db 12773 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 12813
|||||

RESULT 6
HS404G5/c
LOCUS HS404G5 119737 bp DNA linear PRI 04-MAR-2003
DEFINITION Human DNA sequence from clone RP3-404G5 on chromosome 6q24.1-25.2, complete sequence.
ACCESSION AL035695
VERSION AL035695.17 GI:5668650
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 119737)
Wall,M.
Direct Submission
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Jul 31, 1999 this sequence version replaced gi:5650586.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
RP3-404G5 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
Location/Qualifiers
1. .119737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q24.1-25.2"
/clone="RP3-404G5"
/clone_lib="RPCI-3"

ORIGIN

Query Match 96.1%; Score 39.4; DB 9; Length 119737;
Best Local Similarity 97.6%; Pred. No. 0.00031;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 41
|||||
Db 56366 TCTCTCCTGACTTTGTCCGTAATTAAGATCTAATTG 56326
|||||

RESULT 7

AJ627213/c
LOCUS AJ627213 210954 bp DNA linear VRT 14-FEB-2004
DEFINITION Gallus gallus primary C-C chemokine receptor cluster.
ACCESSION AJ627213
VERSION AJ627213.1 GI:42557656
KEYWORDS C-C chemokine receptor 11 like; C-C chemokine receptor 5 like; C-C chemokine receptor 8 like; C-C chemokine receptor 9 like; CCR1-L gene; CCR5-L gene; CCR8-L gene; CCR9-L gene; CCXC chemokine receptor 1 like; CCXCRI-L gene; FCY01 gene; FYVE and coiled-coil. Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE

1
AUTHORS Morroll,S. and Bumstead,N.
TITLE Identification and characterisation of the primary C-C chemokine receptor gene cluster in chicken
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 210954)
AUTHORS Morroll,S.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-2004) Morroll S., Avian Genetics, Institute for Animal Health, Compton Lab., Compton, Newbury, Berks, RG20 7NN, UNITED KINGDOM

FEATURES

Location/Qualifiers
1. .210954
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/chromosome="2"
complement(24094..24960)
source
gene

gene
CDS

```
/proliral  
/mol_type="genomic DNA"  
/isolate="19-108V2"  
/db_xref="taxon:11676"  
/country="USA"  
/note="sample from mother; this mother had an infant who  
initially had a positive PCR test for the HIV virus,  
subsequently the infant seroreverted and had repeated  
negative PCR tests; sequences from the infant's sporadic  
positive PCR test (see GenBank Accession Numbers AF065555  
and AF065556) and from the mother were compared and found  
to be unrelated; Frenkel et al's interpretation of the  
infant's initial, positive PCR test is that it was caused  
by contaminants or other laboratory errors, and the infant  
had not in fact been infected; other researchers have  
interpreted this as reflective of a transient infection"  
<1..>614  
/gene="env"  
<1..>614  
/gene="env"  
/note="C2-V5 region; gp120"  
/codon_start=2  
/product="envelope glycoprotein"  
/protein_id="AAC21517.1"  
/db_xref="GI:3211887"  
/translation="EXEVVIRSENFEDNVKTIIVQLKSVKINCTRPNNNTRKSIPIG  
PGRAYTTGDIIGDIRQAHONISKVEWNNLTKQIVRKLREQFVNKTIIFNQSSGGDPE  
IVLHSEFCGGEFFCYNSTELFNSTWNGNDTEQINNTIGNDTFLPCRIKQIINMWQEV  
GKAMYAPPPIRGXIRCSSNITGLLLRDGGNENGTETFRPGGGD"
```

ORIGIN

Query Match 58.5%; Score 24; DB 14; Length 614;
Best Local Similarity 84.4%; Pred. No. 1.8e+02;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 8 TGACTTTGTCGTAATAATTGATAAGATCTAATT 39
|||||
Db 47 TGACATTGTCGTAATAATTTTCAGATCTAATT 16

RESULT 10
AC010497
LOCUS AC010497 131823 bp DNA linear HTG 04-OCT-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2353F3, WORKING DRAFT SEQUENCE,
3 ordered pieces.
ACCESSION AC010497
VERSION AC010497.6 GI:14579700
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131823)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 131823)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 1, 2001 this sequence version replaced gi:13699537.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 763224
Center clone name: CITB-H1_2353F3

Summary Statistics
Consensus quality: 127285 bases at least Q40

Consensus quality: 129762 bases at least Q30
Consensus quality: 130660 bases at least Q20
Estimated insert size: 108000; pulse field gel estimation
Estimated insert size: 131623; sum-of-contigs estimation
Quality coverage: 7.54 in Q20 bases; pulse field gel estimation
Quality coverage: 6.19 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 25177: contig of 25177 bp in length
* 25178 25277: gap of unknown length
* 25278 52649: contig of 27372 bp in length
* 52650 52749: gap of unknown length
* 52750 131823: contig of 79074 bp in length.
FEATURES
source
1..131823
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2353F3"
/clone_lib="CalTech human BAC library D"

Query Match 58.5%; Score 24; DB 2; Length 131823;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 TCTCTCTGACTTTGTCCGTAATTCGATAAGATCTAATT 40
|||||
Db 68042 TCACTGATAAAGTCTACTCAGTGATATATAAGATCTAATT 68081

RESULT 11
AC010296/c
LOCUS AC010296 134349 bp DNA linear PRI 27-APR-2002
DEFINITION Homo sapiens chromosome 5 clone CTB-155H23, complete sequence.
ACCESSION AC010296
VERSION AC010296.6 GI:20334544
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 134349)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 134349)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 134349)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 134349)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Apr 27, 2002 this sequence version replaced gi:15808509.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center

This Page Blank (uspto)

Search 2

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:44:23 ; Search time 687.273 Seconds
(without alignments)
2566.384 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306168_306503
Perfect score: 336
Sequence: 1 ggatacgaagaccgaaga.....ctggcggaagaggtgccag 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	1044	6	ABZ23390 Nucleotid
2	336	100.0	1044	6	ABZ23391 Reverse c
3	336	100.0	1223	4	AAF29924 Human est
4	336	100.0	1237	6	ABK89697 Oestrogen
5	336	100.0	1374	4	AAF29925 Human est
6	336	100.0	1380	4	AAC86919 Nucleotid
7	336	100.0	1770	10	ADB99351 Fusion pr
8	336	100.0	1785	10	AAD64536 Human wil
9	336	100.0	1788	6	ABL57497 Human nuc
10	336	100.0	1803	10	ADB99349 Fusion pr
11	336	100.0	2028	10	AAD64590 Human ER3
12	336	100.0	2031	10	AAD64588 Human ER3
13	336	100.0	2092	1	AAN70880 cDNA enco
14	336	100.0	2092	4	AAC86920 Nucleotid
15	336	100.0	2092	6	ABK89700 Oestrogen
16	336	100.0	2092	6	ABL51900 Human oes
17	336	100.0	2092	8	AAD53909 Human wil
18	336	100.0	2092	10	ADC09969 Human oes
19	336	100.0	2106	6	ABK89699 Oestrogen
20	336	100.0	2178	10	AAD64589 Human ER3
21	336	100.0	2181	10	AAD64587 Human ER3

22	336	100.0	2220	6	ABS56328 Human oes
23	336	100.0	2220	6	ABA01104 Fused ER
24	336	100.0	2220	10	ADC51161 Human oes
25	336	100.0	2322	2	AAX60628 SSR-LBD f
26	336	100.0	2532	10	ADC51504 Human GST
27	336	100.0	6450	2	AZ223433 Human est
28	336	100.0	6450	6	ABK89695 Oestrogen
29	336	100.0	6450	6	ABK89703 Oestrogen
30	336	100.0	6450	6	ABV94431 Breast ca
31	336	100.0	6450	6	ABZ23389 Reverse c
32	336	100.0	6450	6	ABZ23388 Nucleotid
33	336	100.0	6450	8	ABX12085 cDNA enco
34	336	100.0	6450	8	ACC50140 Breast ca
35	336	100.0	6450	8	ABX93797 Human oes
36	336	100.0	6450	10	ADB81383 DNA seque
37	336	100.0	6450	10	ADD25499 Binding d
38	336	100.0	6450	10	ADE12135 Human oes
39	336	100.0	6450	10	ADF76402 Novel hum
40	336	100.0	6450	10	ADG89355 Cancer de
41	336	100.0	6450	12	ADF83093 Human tes
42	336	100.0	6450	12	ADF72329 Human and
43	336	100.0	6450	12	ADP07306 Human ESR
44	336	100.0	6450	12	ADP05662 Human nuc
45	336	100.0	6468	8	ACC46160 Human dit

ALIGNMENTS

RESULT 1				
ABZ23390				
ID	ABZ23390	standard; cDNA; 1044 BP.		
XX				
AC	ABZ23390;			
XX				
DT	07-APR-2003	(first entry)		
XX				
DE	Nucleotide sequence of human oestrogen receptor alpha bait			
XX				
KW	Human; oestrogen receptor alpha; ER-alpha; CF16; CF17; CF18;			
KW	CF41; CF42; CF43; cofactor; osteoporosis; bone disease; r			
KW	hormonal dysfunction; cancer; cardiovascular disease; ath			
KW	hot flush; mood change; Alzheimer's disease; gene; ss.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	1..1044		
FT		/*tag= a		
FT		/product= "oestrogen receptor alpha bait"		
XX				
PN	WO200270699-A2.			
XX				
PD	12-SEP-2002.			
XX				
PF	28-FEB-2002; 2002WO-EP002189.			
XX				
PR	01-MAR-2001; 2001EP-00105062.			
XX				
PA	(LION-) LION BIOSCIENCE AG.			
XX				
PI	Albers M, Ellwanger S, Loeser E, Koegl M;			
XX				
DR	WPI; 2002-713451/77.			
DR	P-PSDB; ABP70164.			
XX				
PT	New cofactors of estrogen receptor alpha, designated as CF			
PT	CF19, CF40, CF41, CF42 and/or CF43, useful for screening o			
PT	treating osteoporosis, hormonal dysfunctions, cancer or ca			
PT	diseases.			
XX				
PS	Disclosure; Page 108-109; 111pp; English.			
XX				

CC The present sequence encodes a human oestrogen receptor alpha (ER-alpha)
CC bait protein. The specification describes cofactors of ER-alpha,
CC designated CF16, CF17, CF18, CF19, CF40, CF41, CF42, and CF43. The
CC cofactor polypeptides and nucleic acid molecules are useful for screening
CC for compounds for treating osteoporosis and other bone diseases, failures
CC in reproductive functions or hormonal dysfunctions, cancer or
CC cardiovascular diseases such as atherosclerosis, and in preventing hot
CC flushes, mood changes and Alzheimer's disease. The CF proteins are also
CC useful for screening for ligands of the ER alpha. The nucleic acid
CC sequences are useful for making vectors and CF polypeptides, transforming
CC host cells, as research tools for developing nucleic acid probes, and for
CC developing analytical tools such as antisense oligonucleotides
XX
SQ Sequence 1044 BP; 252 A; 272 C; 295 G; 225 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 6; Length 1044;
Best Local Similarity 100.0%; Pred. No. 9.9e-92;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
17 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 76

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
77 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 136

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGCGCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
137 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGCGCGACC 196

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
197 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 256

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGCTTACTGACCAACCTGGCGACAGAGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
257 CCAGACCCCTTCAGTGAAGCTTCGATGATGGCTTACTGACCAACCTGGCGACAGAGGAGC 316

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
317 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 352

RESULT 2
ABZ23391/c
ID ABZ23391 standard; cDNA; 1044 BP.
XX
AC ABZ23391;
XX
DT 07-APR-2003 (first entry)
XX
DE Reverse complement of human oestrogen receptor alpha bait protein cDNA.
XX
KW Human; oestrogen receptor alpha; ER-alpha; CF16; CF17; CF18; CF19; CF40;
KW CF41; CF42; CF43; cofactor; osteoporosis; bone disease; reproduction;
KW hormonal dysfunction; cancer; cardiovascular disease; atherosclerosis;
KW hot flush; mood change; Alzheimer's disease; ss.
XX
OS Homo sapiens.
XX
PN WO200270699-A2.
XX
PD 12-SEP-2002.
XX
PF 28-FEB-2002; 2002WO-EP002189.
XX
PR 01-MAR-2001; 2001EP-00105062.
XX
PA (LION-) LION BIOSCIENCE AG.
XX
PI Albers M, Ellwanger S, Loeser E, Koegl M;
XX

DR WPI; 2002-713451/77.
XX
PT New cofactors of estrogen receptor alpha, designated as CF16, CF17, CF18,
PT CF19, CF40, CF41, CF42 and/or CF43, useful for screening of compounds for
PT treating osteoporosis, hormonal dysfunctions, cancer or cardiovascular
PT diseases.
XX
PS Disclosure; Page 109-110; 111pp; English.
XX
CC The present sequence represents the reverse complement of an oestrogen
CC receptor alpha (ER-alpha) bait protein. The specification describes
CC cofactors of ER-alpha, designated CF16, CF17, CF18, CF19, CF40, CF41,
CC CF42, and CF43. The cofactor polypeptides and nucleic acid molecules are
CC useful for screening for compounds for treating osteoporosis and other
CC bone diseases, failures in reproductive functions or hormonal
CC dysfunctions, cancer or cardiovascular diseases such as atherosclerosis,
CC and in preventing hot flushes, mood changes and Alzheimer's disease. The
CC CF proteins are also useful for screening for ligands of the ER alpha.
CC The nucleic acid sequences are useful for making vectors and CF
CC polypeptides, transforming host cells, as research tools for developing
CC nucleic acid probes, and for developing analytical tools such as
CC antisense oligonucleotides
XX
SQ Sequence 1044 BP; 225 A; 295 C; 272 G; 252 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 6; Length 1044;
Best Local Similarity 100.0%; Pred. No. 9.9e-92;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1028 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 969

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
968 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 909

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGCGCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
908 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGCGCGACC 849

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
848 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 789

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGCTTACTGACCAACCTGGCGACAGAGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
788 CCAGACCCCTTCAGTGAAGCTTCGATGATGGCTTACTGACCAACCTGGCGACAGAGGAGC 729

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
728 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 693

RESULT 3
AAF29924
ID AAF29924 standard; DNA; 1223 BP.
XX
AC AAF29924;
XX
DT 04-APR-2001 (first entry)
XX
DE Human estrogen receptor alpha isoform #1 DNA.
XX
KW Human; estrogen receptor alpha; cancer; osteoporosis; bone; Alzheimer's;
KW cardiovascular; ds.
XX
OS Homo sapiens.
XX
PN WO200100823-A1.
XX
PD 04-JAN-2001.
XX

PF 27-JUN-2000; 2000WO-EP005981.
XX
PR 29-JUN-1999; 99IT-MI001433.
XX
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX
PI Gannon F, Denger S, Flouriot G;
XX
DR WPI; 2001-137955/14.
XX
PT Novel isoforms of human estrogen receptor alpha useful for preparing
PT therapeutic agents for treating cancer, osteoporosis, Alzheimer's disease
PT and cardiovascular diseases.
XX
PS Claim 1; Pag e44; 53pp; English.
XX
CC The present invention relates to a human estrogen receptor (hER)-alpha
CC isoform. Molecules which modulate the activity of the estrogen receptor
CC are useful for the preparation of therapeutic agents for treating cancer,
CC osteoporosis and other bone disorders, Alzheimer's disease and
CC cardiovascular diseases
XX
SQ Sequence 1223 BP; 303 A; 316 C; 341 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 4; Length 1223;
Best Local Similarity 100.0%; Pred. No. 1.1e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCCGAGAGGAGGAGAGATGTTGAAACACACGCGCCAGAGATGATG 60
Db 196 GGATACGAAAAGACCCGAGAGGAGGAGAGATGTTGAAACACACGCGCCAGAGATGATG 255

QY 61 GGGAGGCGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 256 GGGAGGCGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 315

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 316 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 375

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATCTATCTATTCGAGTATGATCCTA 240
Db 376 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATCTATCTATTCGAGTATGATCCTA 435

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC 300
Db 436 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC 495

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 496 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 531

RESULT 4
ABK89697
ID ABK89697 standard; DNA; 1237 BP.
XX
AC ABK89697;
XX
DT 05-NOV-2002 (first entry)
XX
DE Oestrogen receptor alpha nucleic acid comprising A908G mutation, #3.
XX
KW Oestrogen receptor alpha; breast cancer; pre-malignant lesion;
KW invasive breast cancer; A908G oestrogen receptor alpha transition;
KW oestrogen receptor alpha K303R substitution; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200257283-A1.
XX
PD 25-JUL-2002.
XX

PF 16-JAN-2002; 2002WO-US004982.
XX
PR 19-JAN-2001; 2001US-0262990P.
PR 09-JUL-2001; 2001US-0304018P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Fuqua S, O'connell P, Allred DC, Hopp TA;
XX
DR WPI; 2002-590711/63.
XX
PT New isolated estrogen receptor alpha with A908G mutation or K303R
PT substitution, useful as diagnostic marker in breast tissue such as pre-
PT malignant lesions for the development of breast cancer, particularly
PT invasive breast cancer.
XX
PS Claim 1; Page; 133pp; English.
XX
CC The invention relates to an isolated oestrogen receptor alpha nucleic
CC acid sequence comprising an A908G mutation, or an amino acid sequence
CC comprising a K303R substitution. Also described are methods for detecting
CC susceptibility to development of breast cancer or invasive breast cancer
CC in an individual, for diagnosing breast cancer in an individual; and for
CC screening for a modulator of an oestrogen receptor alpha polypeptide
CC comprising a K303R substitution. The oestrogen receptor alpha is useful
CC as a diagnostic marker in breast tissue such as pre-malignant lesions
CC for the development of breast cancer, particularly invasive breast
CC cancer. The methods are useful for determining susceptibility to
CC development of breast cancer, for diagnosing, preventing or treating
CC breast cancer. Transgenic mice may be used for screening and identifying
CC agents that interact with the oestrogen receptor alpha, or affect breast
CC tissue health. The A908G oestrogen receptor alpha transition is
CC frequently present in pre-malignant lesions of the breast and can occur
CC in the adjacent normal-appearing breast epithelium. The present sequence
CC represents an oestrogen receptor alpha nucleic acid which may comprise
CC the A908G mutation. Note: The present sequence is not shown in the
CC specification but was obtained by the indexer from GenBank using the
CC accession number given in the specification
XX
SQ Sequence 1237 BP; 369 A; 218 C; 286 G; 364 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 6; Length 1237;
Best Local Similarity 100.0%; Pred. No. 1.1e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCCGAGAGGAGGAGATGTTGAAACACACGCGCCAGAGATGATG 60
Db 350 GGATACGAAAAGACCCGAGAGGAGGAGATGTTGAAACACACGCGCCAGAGATGATG 409

QY 61 GGGAGGCGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 410 GGGAGGCGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 469

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 470 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 529

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATCTATTCGAGTATGATCCTA 240
Db 530 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATCTATTCGAGTATGATCCTA 589

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC 300
Db 590 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC 649

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 650 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 685

RESULT 5
AAF29925
ID AAF29925 standard; DNA; 1374 BP.

RESULT 7
ADB99351 standard; DNA; 1770 BP.
XX
AC ADB99351;
XX
DT 04-DEC-2003 (first entry)
XX
DE Fusion protein GAL4-AD-ERadAB DNA sequence related to ER constructs.
XX
KW yeast two hybrid assay; oestrogen receptor-alpha receptor;
KW ER-alpha receptor; oestrogen receptor; ER; 17beta-oestradiol;
KW nuclear receptor; steroid receptor; transactivation factor; TAF-1;
KW yeast two hybrid system; GAL4-AD-ERadAB; fusion protein; gene; ds; human;
KW yeast.
XX
OS Chimeric.
OS Unidentified.
OS Homo sapiens.
OS Ascomycota.
XX
FH Key Location/Qualifiers
FT CDS 1..1770
FT /*tag= b
FT /product= "GAL4-AD-ERadAB fusion protein"
FT misc_feature 1..417
FT /*tag= a
FT /note= "This region encodes the GAL4 activation domain
FT from yeast"
FT 418..1770
FT /*tag= c
FT /note= "This region encodes the truncated human ERalpha"
XX
PN WO2003070975-A2.
XX
PD 28-AUG-2003.
XX
PF 21-FEB-2003; 2003WO-EP050029.
XX
PR 25-FEB-2002; 2002EP-00075788.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Verdonk G, Dijkema R, Schoonen WGEJ;
XX
DR WPI; 2003-663859/62.
DR P-PSDB; ADB99352.
XX
PT New gene construct comprising the code for a truncated estrogen receptor
PT (ER)-alpha receptor having regions A and B that are absent, useful for
PT testing compound for their influence of dimerization.
XX
PS Example 3; Page 17-19; 27pp; English.
XX
CC This invention relates to a novel gene construct for use in a yeast two
CC hybrid assay, comprising the code for a truncated oestrogen receptor-
CC alpha (ER-alpha) receptor having regions A and B that are absent. An
CC oestrogen receptor (ER) is a protein that binds the natural hormone
CC 17beta-oestradiol and belongs to the family of nuclear receptors and more
CC specifically to the subfamily of steroid receptors. ER-alpha contains
CC transactivation factor (TAF-1) regions A and B which are absent in the
CC truncated ER-alpha of the current invention. This provides the advantage
CC that the false negative or basal activation in the yeast two hybrid
CC system is reduced in comparison to the full length construct. The gene
CC construct may be useful for use in the yeast two hybrid assay for testing
CC promoting function of dimerisation and/or the transcription
CC compounds which is selectively active on the ER-alpha/ER-beta
CC heterodimer. The present sequence is the DNA sequence which encodes a
CC GAL4-AD-ERadAB fusion protein used in the exemplification of the
CC invention.
XX
SQ Sequence 1770 BP; 489 A; 438 C; 452 G; 391 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 10; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.2e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db |||||
743 GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG 802
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGSCAA 120
Db |||||
803 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGSCAA 862
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCTTGTCCCTGACGCGCGACC 180
Db |||||
863 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCTTGTCCCTGACGCGCGACC 922
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCGCCCCCATCTCTATTCCGAGTATGATCCTA 240
Db |||||
923 AGATGGTCAGTGCCTTGTGGATGCTGAGCGCCCCCATCTCTATTCCGAGTATGATCCTA 982
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db |||||
983 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1042
QY 301 TGGTTCACATGATCAAACTGGGCGAAGAGGGTGCCAG 336
Db |||||
1043 TGGTTCACATGATCAAACTGGGCGAAGAGGGTGCCAG 1078
RESULT 8
AAD64536
ID AAD64536 standard; DNA; 1785 BP.
XX
AC AAD64536;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human wild-type oestrogen receptor (ER) alpha DNA.
KW Tamoxifen activated system; TAS; 4-hydroxytamoxifen; OHT; gene therapy;
KW oestrogen receptor; ER; fundamental research; transgenic; biomedical;
KW disease model; breast cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1785
FT /*tag= a
FT /product= "Human ERalpha protein"
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN US2003199022-A1.
XX
PD 23-OCT-2003.
XX
PF 11-MAR-2002; 2002US-00095373.
XX
PR 11-MAR-2002; 2002US-00095373.
XX
PA (UNII) UNIV ILLINOIS FOUND.
XX
PI Mao C, Shapiro DJ;
XX
DR WPI; 2003-852787/79.
XX
PT New isolated polynucleotide useful for gene therapy applications (e.g. in
PT treating breast cancer), in vivo and in vitro gene expression, in
PT producing bioactive or toxic polypeptides, in research, or in producing
PT transgenic animals.
XX
PS Disclosure; SEQ ID NO 1; Opp; English.

XX The invention relates to tamoxifen activated system (TAS) and 4-
CC hydroxytamoxifen (OHT) activated system for regulated production of
CC proteins in eukaryotic cells. TAS includes mutant oestrogen receptors
CC (ERs) and chimeras thereof. The invention is useful in gene therapy (e.g.
CC in treating breast cancer) or in vivo and in vitro gene expression, in
CC producing bioactive, toxic, recombinant polypeptides in mammalian cells,
CC in biomedical and fundamental research, or in producing transgenic
CC animals. The transgenic animals may be used as disease models, in
CC studying the function and/or activity of a polypeptide, or in identifying
CC and/or evaluating modulators of a polypeptide activity. The present
CC sequence is human wild-type ERalpha DNA used in TAS
XX
SQ Sequence 1785 BP; 417 A; 515 C; 506 G; 347 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 10; Length 1785;
Best Local Similarity 100.0%; Pred. No. 1.2e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
761 GGATACGAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
821 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGGCCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
881 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 9
ABL57497
ID ABL57497 standard; cDNA; 1788 BP.

XX
AC ABL57497;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human nuclear oestrogen receptor alpha coding sequence.
XX
KW Oestrogen; receptor; human; transgenic mouse; cytostatic;
KW antiinflammatory; antidiabetic; endocrine; anorectic; hepatotropic; gene;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1788
FT /*tag= a
FT /product= "Nuclear oestrogen receptor"
FT
XX
PN WO200228175-A2.
XX
PD 11-APR-2002.
XX
XX 28-SEP-2001; 2001WO-IB002246.
PF
XX 03-OCT-2000; 2000FR-00012570.
PR
PR 11-MAY-2001; 2001US-00853033.

XX
PA (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.
XX
XX Chambon P, Metzger D;
XX
DR WPI; 2002-463217/49.
DR P-PSDB; ABB76378.
XX
PT A transgenic mouse, useful in screening for medicaments for the treatment
PT of e.g., diabetes or skin cancers, comprises a fusion protein between a
PT recombinase Cre, and a modified ligand binding domain of the nuclear
PT estrogen receptor alpha.
XX
PS Disclosure; Page 128-130; 149pp; English.
XX
CC The present sequence is the coding sequence of the human nuclear
CC oestrogen receptor alpha (I). The invention relates to a non-human
CC metazoan organism, especially a transgenic mouse, characterised in that
CC at least one cell comprises: (i) a fusion protein formed from a
CC recombinase (Cre), a hinge region (preferably derived from human (I)),
CC and a modified ligand binding domain (LBD) of a nuclear oestrogen
CC receptor, such as (I), its fragment or variant; and (ii) one or more
CC genes or DNA sequences of interest belonging to the genome of the
CC organism, into which one or more recognition sites of the recombinase
CC protein are inserted. The (I) LBD domain is preferably modified by a
CC G521R, G400V, or M543A/L544A mutation. The fusion protein has negligible,
CC or even zero, recombinase activity in the presence of a natural ligand
CC such as oestradiol, but recombinase activity is induced by a small
CC quantity of a synthetic ligand that has antioestrogenic activity, e.g.
CC tamoxifen or 4-hydroxytamoxifen. The metazoan organism or its cells, such
CC as epidermal cells, hepatocytes or adipocytes, are useful in carrying out
CC a spatiotemporally controlled site-specific recombination of a DNA
CC sequence of interest in its natural chromatin environment. It is also
CC used in screening of medicaments for pathological conditions associated
CC with an alteration of the expression and/or function of the DNA sequence
CC of interest, such as skin cancer, inflammation, diabetes, alopecia,
CC obesity, or in promoting hepatic regeneration
XX
SQ Sequence 1788 BP; 418 A; 515 C; 508 G; 347 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 6; Length 1788;
Best Local Similarity 100.0%; Pred. No. 1.2e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
761 GGATACGAAAGACCGAAGAGGAGGGAGATGTTGAAACACAAAGCGCCAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
821 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGGCCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
881 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 10
ADB99349
ID ADB99349 standard; DNA; 1803 BP.
XX

AC	ADB99349;		
XX			
DT	04-DEC-2003	(first entry)	
XX			
DE	Fusion protein GAL4-EDB-ERadAB DNA sequence related to ER constructs.		
XX			
KW	yeast two hybrid assay; oestrogen receptor-alpha receptor;		
KW	ER-alpha receptor; oestrogen receptor; ER; 17beta-oestradiol;		
KW	nuclear receptor; steroid receptor; transactivation factor; TAF-1;		
KW	yeast two hybrid system; GAL4-EDB-ERadAB; fusion protein; gene; ds;		
KW	human; yeast.		
XX			
OS	Chimeric.		
OS	Unidentified.		
OS	Homo sapiens.		
OS	Ascomycota.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1803	
FT		/*tag= b	
FT		/product= "GAL4-DBD-ERadAB fusion protein"	
FT	misc_feature	1..450	
FT		/*tag= a	
FT		/note= "This region encodes the GAL4 DNA binding domain from yeast"	
FT	misc_feature	451..1803	
FT		/*tag= c	
FT		/note= "This region encodes the truncated human ERalpha"	
XX			
PN	WO2003070975-A2.		
XX			
PD	28-AUG-2003.		
XX			
PF	21-FEB-2003; 2003WO-EP050029.		
XX			
PR	25-FEB-2002; 2002EP-00075788.		
XX			
PA	(ALKU) AKZO NOBEL NV.		
XX			
PI	Verdonk G, Dijkema R, Schoonen WGEJ;		
XX			
DR	WPI; 2003-663859/62.		
DR	P-PSDB; ADB99350.		
XX			
PT	New gene construct comprising the code for a truncated estrogen receptor (ER)-alpha receptor having regions A and B that are absent, useful for testing compound for their influence of dimerization.		
PT	Example 2; Page 14-16; 27pp; English.		
PS			
XX			
CC	This invention relates to a novel gene construct for use in a yeast two hybrid assay, comprising the code for a truncated oestrogen receptor-alpha (ER-alpha) receptor having regions A and B that are absent. An oestrogen receptor (ER) is a protein that binds the natural hormone 17beta-oestradiol and belongs to the family of nuclear receptors and more specifically to the subfamily of steroid receptors. ER-alpha contains transactivated ER-alpha of the current invention. This provides the advantage that the false negative or basal activation in the yeast two hybrid system is reduced in comparison to the full length construct. The gene construct may be useful for use in the yeast two hybrid assay for testing compound for their influence of dimerisation and/or the transcription promoting function of dimerisation and selecting a compound out of the compounds which is selectively active on the ER-alpha/ER-beta heterodimer. The present sequence is the DNA sequence which encodes a GAL4-EDB-ERadAB fusion protein used in the exemplification of the invention.		
XX			
SQ	Sequence 1803 BP; 495 A; 432 C; 479 G; 397 T; 0 U; 0 Other;		
	Query Match	100.0%;	Score 336; DB 10; Length 1803;
	Best Local Similarity	100.0%;	Pred. No. 1.2e-91;
	Matches 336; Conservative	0; Mismatches	0; Indels 0; Gaps 0;

QY	1	GGATACGAAAGACCCGAAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGATGATG	60
Db	776	GGATACGAAAGACCCGAAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGATGATG	835
QY	61	GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120
Db	836	GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	895
QY	121	GCCCCTCATGATCAACCGCTCTAAGAAGAACAGCCCTTGTCCCTGACGGCCGACC	180
Db	896	GCCCCTCATGATCAACCGCTCTAAGAAGAACAGCCCTTGTCCCTGACGGCCGACC	955
QY	181	AGATGCTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	240
Db	956	AGATGCTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	1015
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC	300
Db	1016	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGC	1075
QY	301	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	336
Db	1076	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	1111
	RESULT 11		
	AAD64590		
ID	AAD64590	standard; DNA; 2028 BP.	
XX			
AC	AAD64590;		
XX			
DT	12-FEB-2004	(first entry)	
XX			
DE	Human ER3M-VP16-C-79 mutant fusion DNA.		
XX			
KW	Tamoxifen activated system; TAS; 4-hydroxytamoxifen; OHT; gene therapy; oestrogen receptor; ER; fundamental research; transgenic; biomedical; disease model; breast cancer; transactivator; VP16; chimeric; human; mutant; gene; ds.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..2028	
FT		/*tag= a	
FT		/product= "Human ER3M-VP16-C-79 mutant fusion protein"	
FT		/note= "CDS does not include stop codon"	
FT	misc_feature	1792..2028	
FT		/*tag= b	
FT		/note= "Human VP16-C-79 DNA"	
PN	US2003199022-A1.		
XX			
PD	23-OCT-2003.		
XX			
PF	11-MAR-2002; 2002US-00095373.		
XX			
PR	11-MAR-2002; 2002US-00095373.		
XX			
PA	(UNII) UNIV ILLINOIS FOUND.		
XX			
PI	Mao C, Shapiro DJ;		
XX			
DR	WPI; 2003-852787/79.		
DR	P-PSDB; ABW02643.		
XX			
PT	New isolated polynucleotide useful for gene therapy applications (e.g. in treating breast cancer), in vivo and in vitro gene expression, in producing bioactive or toxic polypeptides, in research, or in producing transgenic animals.		

XX 25-MAR-2003 (revised)
DT 24-FEB-1991 (first entry)
XX
DE cDNA encoding human oestrogen receptor protein.
XX
KW Human oestrogen receptor protein; ss cDNA; expression system; assay;
KW steroid receptor protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 293..1280
FT /*tag= a
XX
XX
PN WO8705049-A.
XX
PD 27-AUG-1987.
XX
XX 18-FEB-1987; 87WO-US000341.
PF
XX 20-FEB-1986; 86US-00833829.
PR
XX (CALD) CALIFORNIA BIOTECHNOLOGY INC.
PA
XX Shine J;
PI
XX WPI; 1987-250213/35.
DR P-PSDB; AAP70543.
DR
XX Expression system for vertebrate steroid receptor protein - comprising
PT DNA sequence encoding the protein linked to control sequences in
PT eucaryotic hosts.
PT
XX Disclosure; Fig 1-1 - 1-2; 26pp; English.
PS
XX The cDNA can be used in an expression system to express human oestrogen
CC receptor protein. The coding sequence is operably linked to control
CC sequences compatible with eukaryotic host cells. This method allows
CC expression under conditions which favour appropriate post- translational
CC processing. It produces large amts. of purified protein useful in the
CC design of agonist and antagonist cpds., for the study of the mechanism of
CC action of the steroid binding proteins in general, and for use in
CC diagnostic assays for the proteins or antibodies to them. These assays
CC are important in, eg the diagnosis of tumour sensitivities to steroid
CC metabolism. Suitable host cells are VERO, HeLa and CHO cells. (Updated on
CC 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 2092 BP; 473 A; 607 C; 591 G; 421 T; 0 U; 0 Other;
Query Match 100.0%; Score 336; DB 1; Length 2092;
Best Local Similarity 100.0%; Pred. No. 1.3e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAAGACCGAAGAGAGGAGGAGAGATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1053 GGATACGAAAAGACCGAAGAGAGGAGGAGAGATGTTGAAACACAAAGCGCCAGAGAGATG 1112
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1113 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAGAAGACAGCCTGGCCTTGCCCTGACGGCCGACC 180
Db 1173 GCCCGCTCATGATCAAAACGCTCTAAGAGAAGACAGCCTGGCCTTGCCCTGACGGCCGACC 1232
QY 181 AGATGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 1233 AGATGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA 1292
QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1293 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1352

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGAGGGTGCCAG 336
Db 1353 TGGTTCACATGATCAACTGGGCGAAGAGAGGGTGCCAG 1388

RESULT 14

AAC86920
ID AAC86920 standard; cDNA; 2092 BP.
XX

AC AAC86920;

DT 02-APR-2001 (first entry)
XX

DE Nucleotide sequence of the human oestrogen receptor cDNA.

XX
KW Ribozyme; oestrogen-dependent tumour; cell proliferation; glucocorticoid;
KW DNA-binding domain; oestrogen receptor; cancer treatment; breast cancer;
KW ss.
XX

OS Homo sapiens.

XX WO200074485-A1.
PN

XX 14-DEC-2000.
PD

XX 02-JUN-2000; 2000WO-US015243.
PF

XX 04-JUN-1999; 99US-0137470P.
PR

XX (TEXA) UNIV TEXAS.
PA

XX Roy AK, Lavrovsky Y, Tyagi RK, Song CS, Chatterjee B;
PI

XX WPI; 2001-061633/07.
DR

XX Ribozyme having a high substrate specificity for an mRNA encoding a DNA-
PT binding domain of human estrogen receptor, useful for inhibiting estrogen
PT -dependent tumor cell proliferation, particularly breast cancer.
XX

PS Disclosure; Page 8-9; 49pp; English.

XX The specification describes a ribozyme capable of inhibiting oestrogen-
CC dependent tumour cell proliferation and having a high substrate
CC specificity for an mRNA sequence encoding a DNA-binding domain of human
CC estrogen receptor. The ribozyme is free of endonuclease activity for an
CC mRNA having a DNA binding domain of a glucocorticoid. The oestrogen
CC receptor site-specific ribozymes are useful for cancer treatment and
CC therapies, especially for inhibiting oestrogen-dependent tumour cell
CC proliferation, particularly breast cancer. The present sequence
CC represents the human oestrogen receptor cDNA
XX

SQ Sequence 2092 BP; 473 A; 605 C; 593 G; 421 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 4; Length 2092;
Best Local Similarity 100.0%; Pred. No. 1.3e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGAGGAGGAGAGATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1053 GGATACGAAAAGACCGAAGAGAGGAGGAGAGATGTTGAAACACAAAGCGCCAGAGAGATG 1112

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1113 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAGAAGACAGCCTGGCCTTGCCCTGACGGCCGACC 180
Db 1173 GCCCGCTCATGATCAAAACGCTCTAAGAGAAGACAGCCTGGCCTTGCCCTGACGGCCGACC 1232

QY 181 AGATGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 1233 AGATGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA 1292

QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	300
Db	1293	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	1352
QY	301	TGGTTCACATGATCAACTGGGGCGAAGGGTGCCAG	336
Db	1353	TGGTTCACATGATCAACTGGGGCGAAGGGTGCCAG	1388

RESULT 15
ABK89700
ID ABK89700 standard; DNA; 2092 BP.
XX
AC ABK89700;
XX
DT 05-NOV-2002 (first entry)
XX
DE Oestrogen receptor alpha nucleic acid comprising A908G mutation, #6.
XX
KW Oestrogen receptor alpha; breast cancer; pre-malignant lesion;
KW invasive breast cancer; A908G oestrogen receptor alpha transition;
KW oestrogen receptor alpha K303R substitution; human; gene; ds.

New isolated estrogen receptor alpha with A908G mutation or K303R substitution, useful as diagnostic marker in breast tissue such as pre-malignant lesions for the development of breast cancer, particularly invasive breast cancer.

Claim 1; Page; 133pp; English.

The invention relates to an isolated oestrogen receptor alpha nucleic acid sequence comprising an A908G mutation, or an amino acid sequence comprising a K303R substitution. Also described are methods for detecting susceptibility to development of breast cancer or invasive breast cancer in an individual, for diagnosing breast cancer in an individual; and for screening for a modulator of an oestrogen receptor alpha polypeptide comprising a K303R substitution. The oestrogen receptor alpha is useful as a diagnostic marker in breast tissue such as pre- malignant lesions for the development of breast cancer, particularly invasive breast cancer. The methods are useful for determining susceptibility to development of breast cancer, for diagnosing, preventing or treating breast cancer. Transgenic mice may be used for screening and identifying agents that interact with the oestrogen receptor alpha, or affect breast tissue health. The A908G oestrogen receptor alpha transition is frequently present in pre-malignant lesions of the breast and can occur in the adjacent normal-appearing breast epithelium. The present sequence represents an oestrogen receptor alpha nucleic acid which may comprise the A908G mutation. Note: The present sequence is not shown in the specification but was obtained by the indexer from GenBank using the accession number given in the specification

Sequence 2092 BP: 473 A; 605 C; 593 G; 421 T; 0 U; 0 Other;

Query Match	100.0%;	Score 336;	DB 6;	Length 2092;
Best Local Similarity	100.0%;	Pred. No. 1.3e-91;		

	Matches	336;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAGCGCCACAGAGATGATG	60							
DB	1053	GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAGCGCCACAGAGATGATG	1112							
QY	61	GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120							
DB	1113	GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	1172							
QY	121	GCCCGCTCATGATCAAAACGCTCTAAGAAAGACAGCCTGGCCTTGTCCCTGACGCGCCGACC	180							
DB	1173	GCCCGCTCATGATCAAAACGCTCTAAGAAAGACAGCCTGGCCTTGTCCCTGACGCGCCGACC	1232							
QY	181	AGATGGTCAGTGCCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	240							
DB	1233	AGATGGTCAGTGCCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	1292							
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	300							
DB	1293	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	1352							
QY	301	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	336							
DB	1353	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	1388							

Search completed: November 3, 2004, 06:59:27
Job time : 689.273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 05:58:29 ; Search time 154.335 Seconds
(without alignments)
1547.447 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306168_306503
Perfect score: 336
Sequence: 1 ggatacgaagaccgaaga.....ctgggcgaagaggtgccag 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	2322	3 US-08-564-264-2	Sequence 2, Appli
2	336	100.0	6450	3 US-09-041-886-34	Sequence 34, Appl
3	336	100.0	6450	4 US-08-453-998-1	Sequence 1, Appli
4	336	100.0	6450	4 US-10-027-983-3	Sequence 3, Appli
5	333	99.1	8566	4 US-10-027-983-10	Sequence 10, Appl
6	332.8	99.0	4963	1 US-08-076-726-16	Sequence 16, Appl
7	332.8	99.0	4963	1 US-08-260-452-9	Sequence 9, Appli
8	332.8	99.0	4963	2 US-08-481-970-9	Sequence 9, Appli
9	332.8	99.0	4963	2 US-08-897-719-9	Sequence 9, Appli
10	332.8	99.0	4963	3 US-09-163-269-9	Sequence 9, Appli
11	332.8	99.0	4963	4 US-09-281-674-9	Sequence 9, Appli
12	253.8	75.5	1956	3 US-08-693-940-2	Sequence 2, Appli
13	253.8	75.5	1956	4 US-09-566-660-2	Sequence 2, Appli
14	186.2	55.4	1218	4 US-10-027-983-98	Sequence 98, Appl
15	70.8	21.1	2764	4 US-09-893-666A-1	Sequence 1, Appli
16	67	19.9	1458	2 US-08-836-620A-6	Sequence 6, Appli
17	59	17.6	1251	4 US-09-608-088-2	Sequence 2, Appli
18	59	17.6	1251	4 US-09-711-288-2	Sequence 2, Appli
19	59	17.6	1257	4 US-09-608-088-20	Sequence 20, Appl
20	59	17.6	1257	4 US-09-711-288-20	Sequence 20, Appl
21	59	17.6	1434	4 US-09-608-088-1	Sequence 1, Appli
22	59	17.6	1434	4 US-09-711-288-1	Sequence 1, Appli
23	59	17.6	1460	2 US-08-836-620A-4	Sequence 4, Appli
24	59	17.6	1647	3 US-09-139-617-2	Sequence 2, Appli
25	59	17.6	1647	4 US-09-561-741A-2	Sequence 2, Appli
26	59	17.6	1647	4 US-09-558-795-2	Sequence 2, Appli
27	59	17.6	1898	4 US-09-608-088-24	Sequence 24, Appl

28	59	17.6	1898	4 US-09-711-288-24	Sequence 24, Appl
29	56.6	16.8	2568	2 US-08-836-620A-1	Sequence 1, Appli
30	47	14.0	1615	3 US-09-040-508-1	Sequence 1, Appli
31	47	14.0	1615	3 US-09-500-654-1	Sequence 1, Appli
C 32	38.6	11.5	505	4 US-09-621-976-15639	Sequence 15639, A
C 33	37.8	11.2	144	4 US-10-027-983-97	Sequence 97, Appl
C 34	37.8	11.2	7218	1 US-08-232-463-14	Sequence 14, Appl
35	36.2	10.8	2218	4 US-09-814-915A-4	Sequence 4, Appli
36	36.2	10.8	2539	4 US-09-620-312D-454	Sequence 454, App
37	34.6	10.3	4339	4 US-09-484-970B-164	Sequence 164, App
38	33.2	9.9	3715	3 US-09-041-886-10	Sequence 10, Appl
C 39	32.4	9.6	19056	3 US-09-272-032-8	Sequence 8, Appli
C 40	32.4	9.6	19056	4 US-09-443-218-8	Sequence 8, Appli
C 41	31.2	9.3	91	4 US-09-513-999C-15862	Sequence 15862, A
C 42	31	9.2	676	4 US-09-270-767-1688	Sequence 1688, Ap
C 43	31	9.2	676	4 US-09-270-767-16970	Sequence 16970, A
C 44	31	9.2	702	4 US-09-270-767-27	Sequence 27, Appl
C 45	31	9.2	702	4 US-09-270-767-15309	Sequence 15309, A

ALIGNMENTS

RESULT 1
US-08-564-264-2
; Sequence 2, Application US/08564264
; Patent No. 6040430
; GENERAL INFORMATION:
; APPLICANT: STEWART, Francis
; TITLE OF INVENTION: REGULATION OF SITE-SPECIFIC
; TITLE OF INVENTION: RECOMBINATION BY SITE-SPECIFIC RECOMBINASE/NUCLEAR
; TITLE OF INVENTION: RECEPTOR FUSION PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,264
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: PCT/EP94/02088
; APPLICATION NUMBER: PCT/EP94/02088
; FILING DATE: 28-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 110 298.2
; FILING DATE: 28-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-5019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1-1269)
; OTHER INFORMATION: /note= "FLP recombinase domain."
; FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (1270-1284)
; OTHER INFORMATION: /note= "Linker peptide."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1285-2322)
; OTHER INFORMATION: /note= "Estrogen binding domain."
US-08-564-264-2

Query Match      100.0%; Score 336; DB 3; Length 2322;
Best Local Similarity 100.0%; Pred. No. 2e-99;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db      1295 GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG 1354

Qy      61  GGGAGGGCAGGGGTCAAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db      1355 GGGAGGGCAGGGGTCAAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1414

Qy      121  GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTTGGCCTTGTCCCTGACGCGCCGACC 180
Db      1415 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTTGGCCTTGTCCCTGACGCGCCGACC 1474

Qy      181  AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db      1475 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 1534

Qy      241  CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db      1535 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1594

Qy      301  TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db      1595 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1630

```

```

RESULT 2
US-09-041-886-34
; Sequence 34, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharoz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 361..2146
US-09-041-886-34

Query Match      100.0%; Score 336; DB 3; Length 6450;
Best Local Similarity 100.0%; Pred. No. 3.1e-99;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGATACGAAAAGACCGAAGAGGAGGAGGAGAATGTTGAAACACAAAGCCACAGAGATGATG 60
      |||||||
Db      1121 GGATACGAAAAGACCGAAGAGGAGGAGAATGTTGAAACACAAAGCCACAGAGATGATG 1180

QY      61  GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
      |||||||
Db      1181 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1240

QY      121  GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
      |||||||
Db      1241 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 1300

QY      181  AGATGGTCAGTGCCTTGTGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
      |||||||
Db      1301 AGATGGTCAGTGCCTTGTGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 1360

QY      241  CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
      |||||||
Db      1361 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1420

QY      301  TGGTTCATGATCAACTGGGCGAAGAGGGTGCCAG 336

Db      1421 TGGTTCATGATCAACTGGGCGAAGAGGGTGCCAG 1456

```

RESULT 3
 US-08-453-998-1
 ; Sequence 1, Application US/08453998
 ; Patent No. 644438
 ; GENERAL INFORMATION:
 ; APPLICANT: CHAMON, PIERRE
 ; APPLICANT: METZGER, DANIEL
 ; APPLICANT: WHITE, JOHN
 ; TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
 ; TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND
 ; TITLE OF INVENTION: CORRESPONDING TRANSFORMED STRAINS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVE., N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/453,998
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/161,064
 ; FILING DATE: 03-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CHAPIN, MARLANA K.
 ; REGISTRATION NUMBER: 35,843

```
; REFERENCE/DOCKET NUMBER: 1037/98493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-453-998-1

Query Match          100.0%; Score 336; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 3.1e-99;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1121 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 1180

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1181 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1240

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCCTGTCCCTGACGGCCGACC 180
Db 1241 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCCTGTCCCTGACGGCCGACC 1300

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCTCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1301 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCTCCCATACTCTATTCCGAGTATGATCCTA 1360

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1361 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1420

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1421 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1456

RESULT 4
US-10-027-983-3
; Sequence 3, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(2302)
; NAME/KEY: exon:exon junction
; LOCATION: (1491)...(1492)
; OTHER INFORMATION: exon 4:exon 5
; NAME/KEY: start codon
; LOCATION: (2302)...(2305)
; NAME/KEY: 3'UTR
; LOCATION: (2735)...(8566)
; NAME/KEY: exon:exon junction
; LOCATION: (3569)...(3570)
; OTHER INFORMATION: exon 8:exon 9
; NAME/KEY: exon:exon junction
; LOCATION: (3708)...(3709)
; OTHER INFORMATION: exon 9:exon 10
; NAME/KEY: exon:exon junction
; LOCATION: (3842)...(3843)
; OTHER INFORMATION: exon 10:exon 11
; NAME/KEY: exon:exon junction
; LOCATION: (4026)...(4027)
; OTHER INFORMATION: exon 11:exon 12
US-10-027-983-3

Query Match          100.0%; Score 336; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 3.1e-99;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1121 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 1180

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1121 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 1180

QY 121 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1121 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 1180

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
```

```
Db 1181 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1240
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCCTGTCCCTGACGGCCGACC 180
Db 1241 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCCTGTCCCTGACGGCCGACC 1300
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCTCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1301 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCTCCCATACTCTATTCCGAGTATGATCCTA 1360
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1361 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1420
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1421 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1456

RESULT 5
US-10-027-983-10
; Sequence 10, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 10
; LENGTH: 8566
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(2302)
; NAME/KEY: exon:exon junction
; LOCATION: (1491)...(1492)
; OTHER INFORMATION: exon 4:exon 5
; NAME/KEY: start codon
; LOCATION: (2302)...(2305)
; NAME/KEY: 3'UTR
; LOCATION: (2735)...(8566)
; NAME/KEY: exon:exon junction
; LOCATION: (3569)...(3570)
; OTHER INFORMATION: exon 8:exon 9
; NAME/KEY: exon:exon junction
; LOCATION: (3708)...(3709)
; OTHER INFORMATION: exon 9:exon 10
; NAME/KEY: exon:exon junction
; LOCATION: (3842)...(3843)
; OTHER INFORMATION: exon 10:exon 11
; NAME/KEY: exon:exon junction
; LOCATION: (4026)...(4027)
; OTHER INFORMATION: exon 11:exon 12
US-10-027-983-10

Query Match          99.1%; Score 333; DB 4; Length 8566;
Best Local Similarity 100.0%; Pred. No. 3.4e-98;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 3237 GGATACGAAAGACCGAGAGGGGAGAGTGTGAAACACAAAGCGCCAGAGAGATGATG 3296
QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 3297 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 3356
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTTGTCCCTGACGGCCGACC 180
```


Db 3357 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 3416
QY 181 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db 3417 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 3476
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 3477 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 3536
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGGTGC 333
Db 3537 TGGTTCACATGATCAACTGGGCGAAGAGGGGTGC 3569

RESULT 6
US-08-076-726-16
; Sequence 16, Application US/08076726
; Patent No. 5464758
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Tight Control of Gene Expression in
; TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,726
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-076-726-16

Query Match 99.0%; Score 332.8; DB 1; Length 4963;
Best Local Similarity 99.4%; Pred. No. 3e-98;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAAGAGGAGCGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 1226 GGATACGAAAGACCGAAGAGGAGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 1285
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1286 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1345
QY 121 GCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 180
Db 1346 GCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 1405

QY 181 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db 1406 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1465
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1466 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1525
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGGTGCCAG 336
Db 1526 TGGTTCACATGATCAACTGGGCGAAGAGGGGTGCCAG 1561

RESULT 7
US-08-260-452-9
; Sequence 9, Application US/08260452
; Patent No. 5650298
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochen
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
; TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,452
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: PUHD BGR4
US-08-260-452-9

Query Match 99.0%; Score 332.8; DB 1; Length 4963;
Best Local Similarity 99.4%; Pred. No. 3e-98;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAGCGCCAGAGATGATG 60
Db 1226 GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAGCGCCAGAGATGATG 1285
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120

Db1286GGGAGGCGAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA1345

QY121GCCCCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCCTTGTCCTGACGGCCGACC180

Db1346GCCCCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCCTTGTCCTGACGGCCGACC1405

QY181AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA240

Db1406AGATGGTCATGGCCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA1465

QY241CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC300

Db1466CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC1525

QY301TGGTTACATGATCAACTGGGCGAAGAGGGTGCCAG336

Db1526TGGTTACATGATCAACTGGGCGAAGAGGGTGCCAG1561

RESULT 8

US-08-481-970-9

; Sequence 9, Application US/08481970

; Patent No. 5859310

; GENERAL INFORMATION:

; APPLICANT: Gossen, Manfred

; APPLICANT: Bujard, Hermann

; APPLICANT: Salfeld, Jochen

; APPLICANT: Voss, Jeffrey

; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lahive & Cockfield

; STREET: 60 State Street, Suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/481,970

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/260,452

; FILING DATE: 14-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/076,327

; FILING DATE: 14-JUN-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Giulio A. DeConti, Jr.

; REGISTRATION NUMBER: 31,503

; REFERENCE/DOCKET NUMBER: BBI-013CP2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4963 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Human cytomegalovirus

; IMMEDIATE SOURCE:

; CLONE: pUHD BGR4

US-08-481-970-9

Transcript

Query Match99.0%; Score 332.8; DB 2; Length 4963;

Best Local Similarity99.4%; Pred. No. 3e-98;

Matches 334; Conservative0; Mismatches2; Indels0; Gaps0;

QY1GGATACGAAAGAACCCGAAGAGAGGGAGAGATGTTGAAACACAAGCCGCCAGAGATGATG60

Db1226GGATACGAAAGAACCCGAAGAGAGGGAGAGATGTTGAAACACAAGCCGCCAGAGATGATG1285

QY61GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA120

Db1286GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA1345

QY121GCCCCCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCCTTGTCCTGACGGCCGACC180

Db1346GCCCCCTCATGATCAAAACGCTCTAAGAAGAACAGACCTTGGCCCTTGTCCTGACGGCCGACC1405

QY181AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA240

Db1406AGATGGTCATGGCCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA1465

QY241CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC300

Db1466CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC1525

QY301TGGTTACATGATCAACTGGGCGAAGAGGGTGCCAG336

Db1526TGGTTACATGATCAACTGGGCGAAGAGGGTGCCAG1561

RESULT 9

US-08-897-719-9

; Sequence 9, Application US/08897719

; Patent No. 5922927

; GENERAL INFORMATION:

; APPLICANT: Gossen, Manfred

; APPLICANT: Bujard, Hermann

; APPLICANT: Salfeld, Jochen

; APPLICANT: Voss, Jeffrey

; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic

; TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lahive & Cockfield

; STREET: 60 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/897,719

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/260,452

; FILING DATE:

; APPLICATION NUMBER: 08/076,327

; FILING DATE: 14-JUN-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Giulio A. DeConti, Jr.

; REGISTRATION NUMBER: 31,503

; REFERENCE/DOCKET NUMBER: BBI-013CP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4963 base pairs

; TYPE: nucleic acid


```

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,306
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/260,452
FILING DATE: 14-JUN-1994
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: pUHD BGR4
SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```

	Query Match	99.0%;	Score 332.8;	DB 4;	Length 4963;
	Best Local Similarity	99.4%;	Pred. No. 3e-98;		
	Matches 334;	Conservative	0;	Mismatches	2;
			Indels	0;	Gaps
					0;
QY	1	GGATACGAAAGACCGAAGAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG	60		
Db	1226	GGATACGAAAGACCGAAGAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG	1285		
QY	61	GGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120		
Db	1286	GGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	1345		
QY	121	GCCCGCTCATGATCAACCGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC	180		
Db	1346	GCCCGCTCATGATCAACCGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC	1405		
QY	181	AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCATACTCTATTCCGAGTATGATCCTA	240		
Db	1406	AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCATACTCTATTCCGAGTATGATCCTA	1465		
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	300		
Db	1466	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	1525		
QY	301	TGTTTCAATGATCAACTGGGGCGAAGAGGGTGCCAG	336		
Db	1526	TGTTTCAATGATCAACTGGGGCGAAGAGGGTGCCAG	1561		

RESULT 12
 US-08-693-940-2
 ; Sequence 2, Application US/08693940
 ; Patent No. 6133027
 ; GENERAL INFORMATION:
 ; APPLICANT: Yee, Jiing-Kuan
 ; APPLICANT: Friedman, Theodore
 ; APPLICANT: Chen, Shin-Tai
 ; TITLE OF INVENTION: Inducible Expression System
 ; TITLE OF INVENTION: Useful in the Generation of Packaging Cell Lines for
 ; TITLE OF INVENTION: Pseudotyped Retroviral Vectors
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bozicevic & Reed, LLP
 ; STREET: 285 Hamilton Ave, Suite 200

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,940
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Francis, Carol L
REGISTRATION NUMBER: 36,513
REFERENCE/DOCKET NUMBER: 6510-055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1953
OTHER INFORMATION:
US-08-693-940-2

Query Match 75.5%; Score 253.8; DB 3; Length 1956;
Best Local Similarity 97.4%; Pred. No. 1.1e-72;
Matches 258; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 72 GGTGAAGTGGGCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAAAGCCCGCTCATG 131
Db 1000 GGTGGGATCCATCTGCTGGAGACATGAGAGTGCACACCTTTGGCCAAAGCCCGCTCATG 1059

QY 132 ATCAAACGCTCTAAGAAGAAACAGCCTGGCCTTGTCCTGACGGCCGACAGATGGTCAATG 191
Db 1060 ATCAAACGCTCTAAGAAGAAACAGCCTGGCCTTGTCCTGACGGCCGACAGATGGTCAATG 1119

QY 192 GCCTTGTGGATGCTGAGCCCCCCCCATCTCTATTCCGAGTATGATCCTACAGACCCCTTC 251
Db 1120 GCCTTGTGGATGCTGAGCCCCCCCCATCTCTATTCCGAGTATGATCCTACAGACCCCTTC 1179

QY 252 AGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTACATG 311
Db 1180 AGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTACATG 1239

QY 312 ATCAACTGGGCGAAGAGGTGCCAG 336
Db 1240 ATCAACTGGGCGAAGAGGTGCCAG 1264

```

RESULT 13
US-09-566-660-2
; Sequence 2, Application US/09566660
; Patent No. 6432705
; GENERAL INFORMATION:
; APPLICANT: Yee, Jiing-Kuan
; APPLICANT: Friedmann, Theodore
; APPLICANT: Chen, Shin-Tai
; TITLE OF INVENTION: Inducible Expression System
; FILE REFERENCE: 6510-055CON
; CURRENT APPLICATION NUMBER: US/09/566,660
; CURRENT FILING DATE: 2000-05-08

; PRIOR APPLICATION NUMBER: 08/693,940
; PRIOR FILING DATE: 1996-08-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1953)
; OTHER INFORMATION: Multi-chimeric transactivating factor
US-09-566-660-2

Query Match 75.5%; Score 253.8; DB 4; Length 1956;
Best Local Similarity 97.4%; Pred. No. 1.e-72;
Matches 258; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 72 GGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCCAAGCCCGCTCATG 131
Db 1000 GGTGGGATCCATCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCCAAGCCCGCTCATG 1059
QY 132 ATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCTTGACGGCCGACAGATGGTCAGT 191
Db 1060 ATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCTTGACGGCCGACAGATGGTCAGT 1119
QY 192 GCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTACCAGACCCCTTC 251
Db 1120 GCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTACCAGACCCCTTC 1179
QY 252 AGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGCTGGTTCACATG 311
Db 1180 AGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGAGCTGGTTCACATG 1239
QY 312 ATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1240 ATCAACTGGGCGAAGAGGGTGCCAG 1264

RESULT 14
US-10-027-983-98
; Sequence 98, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 98
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-027-983-98

Query Match 55.4%; Score 186.2; DB 4; Length 1218;
Best Local Similarity 98.4%; Pred. No. 9e-51;
Matches 188; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 146 GAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACAGATGGTCAGTGCCTTGTGGATGC 205
Db 336 GGACCACAGCCTGGCCTTGTCCTTGACGGCCGACAGATGGTCAGTGCCTTGTGGATGC 395
QY 206 TGAGCCCCCCTACTCTATTCCGAGTATGATCCTACCAGACCCCTTCAGTGAAGCTTCGAT 265
Db 396 TGAGCCCCCCTACTCTATTCCGAGTATGATCCTACCAGACCCCTTCAGTGAAGCTTCGAT 455
QY 266 GATGGGCTTACTGACCAACCTGGCAGACAGGAGCTGGTTCACATGATCAACTGGCGAA 325
Db 456 GATGGGCTTACTGACCAACCTGGCAGACAGGAGCTGGTTCACATGATCAACTGGCGAA 515

QY 326 GAGGTGCCAG 336
Db 516 GAGGTGCCAG 526

RESULT 15
US-09-893-666A-1
; Sequence 1, Application US/09893666A
; Patent No. 6759568
; GENERAL INFORMATION:
; APPLICANT: YAMASHITA, ICHIRO
; TITLE OF INVENTION: High estrogen-sensitive medaka fish
; FILE REFERENCE: 210217US-620-7249-0
; CURRENT APPLICATION NUMBER: US/09/893,666A
; CURRENT FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: JP 2000-247729
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2764
; TYPE: DNA
; ORGANISM: Oryzias latipes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (211)..(1935)
; OTHER INFORMATION:
US-09-893-666A-1

Query Match 21.1%; Score 70.8; DB 4; Length 2764;
Best Local Similarity 64.8%; Pred. No. 5e-13;
Matches 105; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 175 CCGACCAGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCGGAGTATG 234
Db 1022 CTGAGCAGGTGCTGCTCCTCTCAGGGCCGAGCCCCCGATACTCTGCTCGGTCAGA 1081
QY 235 ATCCTACCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGCA 294
Db 1082 AGTTGAGCCGACCGTACACCCAGGTCACCATGATGACCCCTGCTCACCAGCATGGCAGACA 1141
QY 295 GGGAGCTGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1142 AGGAGCTGGTCCACATGATCGCCTGGGCCAAGAAGCTCCAG 1183

Search completed: November 3, 2004, 09:46:54
Job time : 156.335 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 06:51:04 ; Search time 611.713 Seconds
(without alignments)
2816.475 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306168_306503
Perfect score: 336
Sequence: 1 ggatacgaagaccgaaga.....ctggggaagagggtgccag 336

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	1237	14	US-10-052-092-3
2	336	100.0	1237	15	US-10-437-107-3
3	336	100.0	1785	15	US-10-095-373A-1
4	336	100.0	1788	9	US-09-853-033-1
5	336	100.0	2028	15	US-10-095-373A-72
6	336	100.0	2031	15	US-10-095-373A-68
7	336	100.0	2092	10	US-09-952-680A-9
8	336	100.0	2092	14	US-10-052-092-6
9	336	100.0	2092	15	US-10-157-899A-54
10	336	100.0	2092	15	US-10-437-107-6
11	336	100.0	2106	14	US-10-052-092-5
12	336	100.0	2106	15	US-10-437-107-5
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 72, Appl
					Sequence 68, Appl
					Sequence 9, Appli
					Sequence 6, Appli
					Sequence 54, Appl
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 5, Appli

13	336	100.0	2178	15	US-10-095-373A-70	Sequence 70, Appl
14	336	100.0	2181	15	US-10-095-373A-66	Sequence 66, Appl
15	336	100.0	6450	13	US-10-096-710-2	Sequence 2, Appli
16	336	100.0	6450	13	US-10-081-563-1	Sequence 1, Appli
17	336	100.0	6450	14	US-10-052-092-1	Sequence 1, Appli
18	336	100.0	6450	14	US-10-052-092-7	Sequence 7, Appli
19	336	100.0	6450	15	US-10-207-655-60	Sequence 60, Appl
20	336	100.0	6450	15	US-10-177-293-127	Sequence 127, App
21	336	100.0	6450	15	US-10-027-983-3	Sequence 3, Appli
22	336	100.0	6450	15	US-10-007-926A-422	Sequence 422, App
23	336	100.0	6450	15	US-10-437-107-1	Sequence 1, Appli
24	336	100.0	6450	15	US-10-437-107-7	Sequence 7, Appli
25	336	100.0	6450	15	US-10-392-274-1	Sequence 1, Appli
26	336	100.0	6450	15	US-10-448-753-3	Sequence 3, Appli
27	336	100.0	6450	15	US-10-172-118-435	Sequence 435, App
28	336	100.0	6450	15	US-10-388-360-303	Sequence 303, App
29	336	100.0	6450	16	US-10-342-887-435	Sequence 435, App
30	336	100.0	6610	14	US-10-052-092-20	Sequence 20, Appl
31	336	100.0	6610	15	US-10-437-107-20	Sequence 20, Appl
32	336	100.0	6833	15	US-10-422-934-10	Sequence 10, Appl
33	336	100.0	6956	15	US-10-422-934-9	Sequence 9, Appli
34	336	100.0	7038	15	US-10-422-934-3	Sequence 3, Appli
35	336	100.0	7979	18	US-10-475-433-9	Sequence 9, Appli
36	336	100.0	11365	17	US-10-433-311A-17	Sequence 17, Appl
37	336	100.0	465237	9	US-09-933-267A-1	Sequence 1, Appli
38	333	99.1	8566	15	US-10-027-983-10	Sequence 10, Appl
39	333	99.1	8566	15	US-10-448-753-10	Sequence 10, Appl
40	332.8	99.0	4963	9	US-09-281-674-9	Sequence 9, Appli
41	332.8	99.0	4963	9	US-09-777-317-9	Sequence 9, Appli
42	332.8	99.0	4963	9	US-09-892-227-9	Sequence 9, Appli
43	325	96.7	1317	15	US-10-157-899A-3	Sequence 3, Appli
44	325	96.7	1317	15	US-10-157-899A-11	Sequence 11, Appl
45	325	96.7	1317	15	US-10-157-899A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-10-052-092-3
; Sequence 3, Application US/10052092
; Publication No. US20030027778A1

; GENERAL INFORMATION:

; APPLICANT: Fuqua, Suzanne

; APPLICANT: Allred, D.

; APPLICANT: Hopp, Torsten A.

; APPLICANT: O'Connell, Peter

; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic

; FILE REFERENCE: P02102US2

; CURRENT APPLICATION NUMBER: US/10/052,092

; CURRENT FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US 60/262,990

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: US 60/304,018

; PRIOR FILING DATE: 2001-07-09

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 1237

; TYPE: DNA

; ORGANISM: Human

US-10-052-092-3

Query Match 100.0%; Score 336; DB 14; Length 1237;

Best Local Similarity 100.0%; Pred. No. 3.1e-104;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAACACACAAAGCGCAGAGATGATG 60

Db 350 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAACACACAAAGCGCAGAGATGATG 409

QY 61 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120

Db 410 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 469

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 180

Db 470 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 529

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240

Db 530 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 589

QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300

Db 590 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 649

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336

Db 650 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 685

RESULT 2

US-10-437-107-3

; Sequence 3, Application US/10437107

; Publication No. US20030186313A1

; GENERAL INFORMATION:

; APPLICANT: Fuqua, Suzanne

; APPLICANT: Allred, D.

; APPLICANT: Hopp, Torsten A.

; APPLICANT: O'Connell, Peter

; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti

; FILE REFERENCE: P02102US2

; CURRENT APPLICATION NUMBER: US/10/437,107

; CURRENT FILING DATE: 2003-05-13

; PRIOR APPLICATION NUMBER: US/10/052,092

; PRIOR FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US 60/262,990

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: US 60/304,018

; PRIOR FILING DATE: 2001-07-09

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 1237

; TYPE: DNA

; ORGANISM: Human

US-10-437-107-3

Query Match 100.0%; Score 336; DB 15; Length 1237;

Best Local Similarity 100.0%; Pred. No. 3.1e-104;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAAGACCGAAGAGGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGATGATG 60

Db 350 GGATACGAAAAAGACCGAAGAGGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGATGATG 409

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120

Db 410 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 469

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 180

Db 470 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 529

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240

Db 530 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 589

QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300

Db 590 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 649

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336

Db 650 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 685

RESULT 3

US-10-095-373A-1

; Sequence 1, Application US/10095373A

; Publication No. US20030199022A1

; GENERAL INFORMATION:

; APPLICANT: Shapiro, David J.

; APPLICANT: Mao, Chengjian

; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULATAE

; TITLE OF INVENTION: PRODUCTION OF PROTEINS IN EUKARYOTIC CELLS

; FILE REFERENCE: 10322.25

; CURRENT APPLICATION NUMBER: US/10/095,373A

; CURRENT FILING DATE: 2002-06-21

; NUMBER OF SEQ ID NOS: 83

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1785

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-095-373A-1

Query Match 100.0%; Score 336; DB 15; Length 1785;

Best Local Similarity 100.0%; Pred. No. 3.4e-104;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAAGACCGAAGAGGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGATGATG 60

Db 761 GGATACGAAAAAGACCGAAGAGGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGATGATG 820

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120

Db 821 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 180

Db 881 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240

Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300

Db 1001 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336

Db 1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 4

US-09-853-033-1

; Sequence 1, Application US/09853033

; Patent No. US20020100068A1

; GENERAL INFORMATION:

; APPLICANT: CHAMBON, PIERRE

; APPLICANT: METZGER, DANIEL

; TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION

; TITLE OF INVENTION: MEDIATED BY MODIFIED CRE-ER

; FILE REFERENCE: 065691/0222

; CURRENT APPLICATION NUMBER: US/09/853,033

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: FR 00/12570

; PRIOR FILING DATE: 2000-10-03

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1788

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

```

; LOCATION: (1)..(1788)
US-09-853-033-1
    Query Match          100.0%; Score 336; DB 9; Length 1788;
    Best Local Similarity 100.0%; Pred. No. 3.4e-104;
    Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db 761 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 5
US-10-095-373A-72
; Sequence 72, Application US/10095373A
; Publication No. US20030199022A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, David J.
; APPLICANT: Mao, Chengjian
; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULAT
; FILE REFERENCE: 10322.25
; CURRENT APPLICATION NUMBER: US/10/095,373A
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: construct
US-10-095-373A-72
    Query Match          100.0%; Score 336; DB 15; Length 2028;
    Best Local Similarity 100.0%; Pred. No. 3.5e-104;
    Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 761 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000
```

```

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 6
US-10-095-373A-68
; Sequence 68, Application US/10095373A
; Publication No. US20030199022A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, David J.
; APPLICANT: Mao, Chengjian
; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULAT
; FILE REFERENCE: 10322.25
; CURRENT APPLICATION NUMBER: US/10/095,373A
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: construct
US-10-095-373A-68
    Query Match          100.0%; Score 336; DB 15; Length 2031;
    Best Local Similarity 100.0%; Pred. No. 3.5e-104;
    Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG 60
Db 1007 GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG 1066

QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1067 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1126

QY 121 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 180
Db 1127 GCCCGCTCATGATCAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGGCCGACC 1186

QY 181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240
Db 1187 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1246

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1247 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1306

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1307 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1342

RESULT 7
US-09-952-680A-9
; Sequence 9, Application US/09952680A
; Publication No. US20030087239A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Marty
; APPLICANT: Epstein, David
; APPLICANT: Hamaguchi, No. US20030087239Aluko
; TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
; FILE REFERENCE: 23239-501
; CURRENT APPLICATION NUMBER: US/09/952,680A
; CURRENT FILING DATE: 2001-09-13
```

```

; PRIOR APPLICATION NUMBER: 60/232,454
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-952-680A-9

Query Match      100.0%; Score 336; DB 10; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGATACGAAAAGACCGAAGAGGAGGAGGAGAGATGTTGAAACACAAAGCGCCGAGAGATGATG 60
      |||
Db      1053 GGATACGAAAAGACCGAAGAGGAGGAGAGATGTTGAAACACAAAGCGCCGAGAGATGATG 1112

QY      61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
      |||
Db      1113 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172

QY      121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTTGCCCTTGCCCTGAGCGCGGACC 180
      |||
Db      1173 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTTGCCCTTGCCCTGAGCGCGGACC 1232

QY      181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
      |||
Db      1233 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1292

QY      241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCGACAGGGGAGC 300
      |||
Db      1293 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCGACAGGGGAGC 1352

QY      301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
      |||
Db      1353 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1388

RESULT 8
US-10-052-092-6
; Sequence 6, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-6

Query Match      100.0%; Score 336; DB 14; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGATACGAAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCGAGAGATGATG 60
      |||
Db      1053 GGATACGAAAAGACCGAAGAGGAGGAGGAGATGTTGAAACACAAAGCGCCGAGAGATGATG 1112

QY      61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
```

```

      |||
Db      1113 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172

QY      121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTTGCCCTTGCCCTGAGCGCGGACC 180
      |||
Db      1173 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCGCTTGCCCTTGCCCTGAGCGCGGACC 1232

QY      181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
      |||
Db      1233 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1292

QY      241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCGACAGGGGAGC 300
      |||
Db      1293 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCGACAGGGGAGC 1352

QY      301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
      |||
Db      1353 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1388

RESULT 9
US-10-157-899A-54
; Sequence 54, Application US/10157899A
; Publication No. US20030143559A1
; GENERAL INFORMATION:
; APPLICANT: Bracken, Kathryn Rene
; APPLICANT: de los Angeles, Joseph Ernest
; APPLICANT: Huang, Ying
; APPLICANT: Kadan, Michael Joseph
; APPLICANT: Ksander, Gary Michael
; APPLICANT: Zerby, Dennis
; TITLE OF INVENTION: NOVEL ESTROGEN RECEPTOR LIGAND BINDING DOMAIN VARIANTS AND NOVEL
; TITLE OF INVENTION: LIGANDS AND PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: 4-32018A
; CURRENT APPLICATION NUMBER: US/10/157,899A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/294,839
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (293)..(2080)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (293)..(832)
; OTHER INFORMATION: A/B domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (833)..(1081)
; OTHER INFORMATION: C domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1082)..(1198)
; OTHER INFORMATION: D domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1199)..(1954)
; OTHER INFORMATION: E domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1955)..(2077)
; OTHER INFORMATION: F domain
;
US-10-157-899A-54

Query Match      100.0%; Score 336; DB 15; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 1053 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 1112
QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 120
Db 1113 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 1172
QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 180
Db 1173 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 1232
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1233 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1292
QY 241 CCAGACCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1293 CCAGACCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1352
QY 301 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 336
Db 1353 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 1388
RESULT 10
US-10-437-107-6
; Sequence 6, Application US/10437107
; Publication No. US20030186313A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/437,107
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US/10/052,092
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Human
US-10-437-107-6
Query Match 100.0%; Score 336; DB 15; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 1053 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 1112
QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 120
Db 1113 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 1172
QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 180
Db 1173 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 1232
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1233 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1292

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1293 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1352
QY 301 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 336
Db 1353 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 1388
RESULT 11
US-10-052-092-5
; Sequence 5, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-5
Query Match 100.0%; Score 336; DB 14; Length 2106;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 761 GGATACGAAAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 820
QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGTGCCAACTTTGGCCAA 880
QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 180
Db 881 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTTGACGGCCGACC 940
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1000
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060
QY 301 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAAACTGGCGGAAGAGGGTGCCAG 1096
RESULT 12
US-10-437-107-5
; Sequence 5, Application US/10437107
; Publication No. US20030186313A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutic


```

; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/437,107
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US/10/052,092
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Human
US-10-437-107-5

Query Match      100.0%; Score 336; DB 15; Length 2106;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 60
Db 761 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1096

RESULT 13
US-10-095-373A-70
; Sequence 70, Application US/10095373A
; Publication No. US20030199022A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, David J.
; APPLICANT: Mao, Chengjian
; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULAT
; TITLE OF INVENTION: PRODUCTION OF PROTEINS IN EUKARYOTIC CELLS
; FILE REFERENCE: 10322.25
; CURRENT APPLICATION NUMBER: US/10/095,373A
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 2178
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: construct
US-10-095-373A-70

Query Match      100.0%; Score 336; DB 15; Length 2178;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 60
Db 761 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1096

RESULT 14
US-10-095-373A-66
; Sequence 66, Application US/10095373A
; Publication No. US20030199022A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, David J.
; APPLICANT: Mao, Chengjian
; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULAT
; TITLE OF INVENTION: PRODUCTION OF PROTEINS IN EUKARYOTIC CELLS
; FILE REFERENCE: 10322.25
; CURRENT APPLICATION NUMBER: US/10/095,373A
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: construct
US-10-095-373A-66

Query Match      100.0%; Score 336; DB 15; Length 2181;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 60
Db 1157 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 1216

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1217 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1276

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 1277 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 1336

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 1337 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 1396

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1397 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1456

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1457 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1492
```

```

Db 761 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 820
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880
QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 940
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 1000
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060
QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1096

RESULT 14
US-10-095-373A-66
; Sequence 66, Application US/10095373A
; Publication No. US20030199022A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, David J.
; APPLICANT: Mao, Chengjian
; TITLE OF INVENTION: A TAMOXIFEN AND 4-HYDROXY TAMOXIFEN-ACTIVATED SYSTEM FOR REGULAT
; TITLE OF INVENTION: PRODUCTION OF PROTEINS IN EUKARYOTIC CELLS
; FILE REFERENCE: 10322.25
; CURRENT APPLICATION NUMBER: US/10/095,373A
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: construct
US-10-095-373A-66

Query Match      100.0%; Score 336; DB 15; Length 2181;
Best Local Similarity 100.0%; Pred. No. 3.6e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 60
Db 1157 GGATACGAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCCAGAGAGATGATG 1216

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1217 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1276

QY 121 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 1277 GCGCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 1336

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 1337 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCGAGTATGATCCTA 1396

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1397 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1456

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1457 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1492
```



```
RESULT 15
US-10-096-710-2
; Sequence 2, Application US/10096710
; Publication No. US20020164581A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Chawnshang
; TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS THAT
; TITLE OF INVENTION: DO NOT ATTENUATE THE PROTECTIVE EFFECTS OF ESTROGEN
; FILE REFERENCE: 21108.0010U2
; CURRENT APPLICATION NUMBER: US/10/096,710
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,229
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: synthetic construct
US-10-096-710-2

Query Match      100.0%; Score 336; DB 13; Length 6450;
Best Local Similarity 100.0%; Pred. No. 5e-104;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGATACGAAAGACCGAGAGAGGGGAGAATGTTGAACACAAAGCGCCAGAGAGATGATG 60
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1121 GGATACGAAAGACCGAGAGAGGGGAGAATGTTGAACACAAAGCGCCAGAGATGATG 1180

QY      61  GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1181 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1240

QY      121 GCCCGCTCATGATCAACCGTCTTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 180
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1241 GCCCGCTCATGATCAACCGTCTTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 1300

QY      181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1301 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 1360

QY      241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1361 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1420

QY      301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
      |||||||||||||||||||||||||||||||||||
Db      1421 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1456
```

Search completed: November 3, 2004, 10:01:04
Job time : 612.713 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:57:13 ; Search time 5800.42 Seconds
(without alignments)
2110.838 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306168_306503
Perfect score: 336
Sequence: 1 ggatacgaagaccgaaga.....ctgggggaagagggtgccag 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hlc:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	611	5	BX489897 DKFZp686N
2	334.4	99.5	653	1	AI073549 OV45e08.x
3	284.8	84.8	705	5	BP160291 BP160291
4	257.6	76.7	616	4	BG084311 H3099A08-
5	257.6	76.7	2109	3	AK039911 Mus muscu
6	257.6	76.7	2591	3	AK077236 Mus muscu
7	257.6	76.7	2735	3	AK054182 Mus muscu
8	257.6	76.7	4248	3	AK041525 Mus muscu
9	257.6	76.7	4315	3	AK087638 Mus muscu
10	257.6	76.7	4321	3	AK036627 Mus muscu
11	245.6	73.1	593	1	AA023625 mh80b06.r
12	222.2	66.1	365	1	AA092184 1L4609.se
13	202.8	60.4	627	2	BB200714 BB200714
14	170.8	50.8	576	1	AU041270 AU041270
15	118.4	35.2	316	2	BB097271 BB097271
16	91.8	27.3	648	2	BF660250 maa95g12.
17	80.4	23.9	459	8	BZ157672 CH230-433
18	68.4	20.4	1203	9	CL025390 CH216-21J
19	66.8	19.9	1277	9	CL025388 CH216-21J
20	66.6	19.8	642	2	BF058527 7k3lh08.x
21	65.4	19.5	2712	3	AK054290 Mus muscu
22	65.4	19.5	3144	3	AK054301 Mus muscu
23	65.4	19.5	3363	3	AK054413 Mus muscu
24	64.4	19.2	1160	6	CD508522 CDA91-C11

C	25	63.2	18.8	906	9	CNS01WT6	AL170835 Tetraodon
	26	61.8	18.4	641	2	AW976636	AW976636 EST388745
	27	60.4	18.0	735	8	BZ834605	BZ834605 CH240_209
	28	59.6	17.7	765	9	CNS03FXW	AL242285 Tetraodon
	29	58.6	17.4	642	2	BF733032	BF733032 nael8e05.
	30	58	17.3	992	9	CNS038DM	AL232483 Tetraodon
C	31	55.8	16.6	955	9	CNS04QE6	AL302487 Tetraodon
	32	53	15.8	422	6	CA383629	CA383629 663832 NC
	33	52	15.5	275	7	CN980443	CN980443 48400_126
	34	48.6	14.5	971	9	CNS03ZWK	AL268157 Tetraodon
C	35	47	14.0	303	4	BG217052	BG217052 RST36753
	36	47	14.0	696	4	BM723962	BM723962 UI-B-E01-
	37	47	14.0	1013	7	CN647617	CN647617 ILLUMIGEN
	38	47	14.0	1377	9	AY412347	AY412347 Homo sapi
	39	47	14.0	1377	9	AY412348	AY412348 Pan trogl
C	40	46.8	13.9	714	9	CNS01VNY	AL169351 Tetraodon
	41	46.4	13.8	979	9	CNS04DMC	AL285933 Tetraodon
C	42	45.4	13.5	312	4	BG211665	BG211665 RST31233
C	43	45.4	13.5	548	7	CO627140	CO627140 DG9-283d1
C	44	44.4	13.2	366	4	BG194599	BG194599 RST13761
	45	44	13.1	1273	4	BI079710	BI079710 602876202

ALIGNMENTS

RESULT 1
BX489897
LOCUS BX489897 611 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686N0562_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION BX489897
VERSION BX489897.1 GI:31998955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 611)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686N0562) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N0562"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 100.0%; Score 336; DB 5; Length 611;
Best Local Similarity 100.0%; Pred. No. 4.8e-84;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAGAGGAGGAGATGTTGAAACACAGCGCCAGAGATGATG 60

Db	5	GGATACGAAAGACCGAAGAGAGGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG	64
QY	61	GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAAACCTTTGGCCAA	120
Db	65	GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAAACCTTTGGCCAA	124
QY	121	GCCCGCTCATGATCAAAACGCTCTAAGAAAGAACAGCCCTGGCCCTTGTCCTGACGGCCGACC	180
Db	125	GCCCGCTCATGATCAAAACGCTCTAAGAAAGAACAGCCCTGGCCCTTGTCCTGACGGCCGACC	184
QY	181	AGATGGTCAGTGCCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA	240
Db	185	AGATGGTCAGTGCCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCGAGTATGATCCTA	244
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	300
Db	245	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	304
QY	301	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	336
Db	305	TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG	340

RESULT 2	AI073549/c	LOCUS	AI073549	653 bp	mRNA	linear	EST 06-AUG-1998
DEFINITION	ov45e08.x1 Soares testis NHT Homo sapiens		ov45e08.x1 Soares testis NHT Homo sapiens		cdna clone		IMAGE:1640294
	3' similar to gb:M12674		ESTROGEN RECEPTOR (HUMAN);		mRNA sequence.		
ACCESSION	AI073549		AI073549				
VERSION	AI073549.1		GI:3400193				
KEYWORDS	EST.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 653)						
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .						
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),						
	Tumor Gene Index						
JOURNAL	Unpublished (1997)						
COMMENT	Contact: Robert Strausberg, Ph.D.						
	Email: cgapbs-r@mail.nih.gov						
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima						
	Bonaldo, Ph.D.						
	CDNA Library Arrayed by: Greg Lennon, Ph.D.						
	DNA Sequencing by: Washington University Genome Sequencing Center						
	Clone distribution: NCI-CGAP clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LLNL at:						
	www-bio.llnl.gov/bbrp/image/image.html						
	Seq primer: -40m13 fwd. ET from Amersham						
	High quality sequence stop: 483.						

	Query Match	99.5%;	Score 334.4;	DB 1;	Length 653;
	Best Local Similarity	99.7%;	Pred. No. 1.4e-83;		
	Matches 335;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps	0;
QY	1	GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG	60		
Db	652	GGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAAGCGCCAGAGATGATG	593		
QY	61	GGGAGGGCAGGGGTGAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120		
Db	592	GGGAGGGCAGGGGTGAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	533		
QY	121	GCCCGCTCATGATCAAAACGCTCTAAGAAGAAACAGCCTGGCCTTGTCCCTGACGGCCGACC	180		
Db	532	GCCCGCTCATGATCAAAACGCTCTAAGAAGAAACAGCCTGGCCTTGTCCCTGACGGCCGACC	473		
QY	181	AGATGGTCAGTGCCTTGTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	240		
Db	472	AGATGGTCAGTGCCTTGTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA	413		
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGTACCAACCTGGCAGACAGGGAGC	300		
Db	412	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGTACCAACCTGGCAGACAGGGAGC	353		
QY	301	TGGTTCACATGATCAACTGGGCGAAGAGGGTGGCCAG	336		
Db	352	TGGTTCACATGATCAACTGGGCGAAGAGGGTGGCCAG	317		

RESULT 3	ACCESSION	REFERENCE
BPI60291	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

FEATURES source

ORIGIN

ORIGIN

Query Match	84.8%	Score 284.8	DB 5	Length 705
Best Local Similarity	90.5%	Pred. No. 1.5e-69		
Matches 304	Conservative 0	Mismatches 32	Indels 0	Gaps 0
QY 1	GGATACGAAAGACCGAAGAGGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60			
Db 359	GGATACGGAAAGACCGGAGAGGAGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 418			
QY 61	GGGAGGGCAGGGGTGAAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120			
Db 419	GAGAGGGCAGGGAATGAAGCGGTGCCCTGGAGACATGAGATCTGCCAACCTTTGGCCAA 478			
QY 121	GCCCGCTCATGATCAAACGCTCTAAGAAGAACAGCGCTGGCCCTTGTCCTGACGGCCGACC 180			
Db 479	GCCCTCTCTTGATTAAACACACTAAGAAGAACAGCGCCGCTTGTCCCTGACAGCCGACC 538			
QY 181	AGATGGTCAGTGCCCTTGTTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240			
Db 539	AGATGATCAGTGCCCTTGTTGGAGGCTGAGCCCCCATAAATCTATTCCGAGTATGATCCTA 598			
QY 241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300			
Db 599	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTCAATGATGGGCTTGCTGACCAACCTCGCAGACAGGGAGC 658			
QY 301	TGGTTTCATGATCAACTGGGGCGAAGAGGGTGCCAG 336			
Db 659	TGGCACACATGATCCACTGGGCAAAAGAGGGTGCCAG 694			

RESULT 4

LOCUS	616 bp	linear	EST 18-DEC-2003
DEFINITION	H3099A08-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone		
ACCESSION	H3099A08 5', mRNA sequence.		
VERSION	BG084311		
KEYWORDS	BG084311.2 GI:40071640		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 616)		
TITLE	Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H.		
JOURNAL	Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)		
PUBMED	20381348		
COMMENT	10922068		
	On Jan 26, 2001 this sequence version replaced gi:12566875.		

FEATURES

Source

/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match	76.7%;	Score 257.6;	DB 4;	Length 616;
Best Local Similarity	85.4%;	Pred. No. 7.3e-62;		
Matches 287;	Conservative	0;	Mismatches 49;	Indels 0; Gaps 0;
QY	1	GGATACGAAAAGACCGAAGAGGAGGGAGAAATGTTGAACACAAAGCGCCAGAGAGATGATG	60	
Db	177	GCATACGAAAAGACCGCCGAGGAGGGAGAAATGTTGAAGCACAAAGCGTCAGAGAGATGACT	236	
QY	61	GGGAGGCGAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120	
Db	237	TGGAAGGCGGAAATGAAATGGTGGCTTCAGGAGACATGAGGGCTGCCAACCTTTGGCCAA	296	
QY	121	GCCCGTCATGATCAAAAGCTCTAAGAAACACAGCCTTGGCCTTGTCCTTGACGCGCGACC	180	
Db	297	GCCCTCTTGATTAAGCACACTAAGAAAGATAGCCCTGCCTTGTCCTTGACAGCTGACC	356	
QY	181	AGATGGTCAGTGCCTTGTGGATGCTGAGCGCCCCCATACTCTATTCCGAGTATGATCCTA	240	
Db	357	AGATGGTCAGTGCCTTGTGGATGCTGAACCGCCCCATGATCTATTCTGAATATGATCCTT	416	
QY	241	CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC	300	
Db	417	CTAGACCCCTTCAGTGAAGCCTCAATGATGGGCTTATTGACCAACCTAGCAGATAGGGAGC	476	
QY	301	TGGTTCACATGATCAACTGGGGCGAAAGAGGGTGCCAG	336	
Db	477	TGGTTCATATGATCAACTGGGCAAGAGAGAGTGCCAG	512	

RESULT

AK039911
LOCUS
DEFINITION
AK039911 2109 bp mRNA linear HTC 03-APR-2004
Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched
library, clone:A430027L01 product:estrogen receptor 1 (alpha), full
insert sequence.
AK039911
ACCESSION
VERSION AK039911.1 GI:26087489
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253

PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. .2109
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:A430027L01"
/db_xref="taxon:10090"
/clone="A430027L01"
/tissue type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"

misc_feature 1. .2109
/note="estrogen receptor 1 (alpha) (MGD|MGI:1352467, GB|NM_007956, evidence: BLASTN, 99%, match=1960)"
ORIGIN
Query Match 76.7%; Score 257.6; DB 3; Length 2109;
Best Local Similarity 85.4%; Pred. No. 9.3e-62;
Matches 287; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAGAGGAGGAGAAATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1458 GCATACGGAAAGACCGCCGAGGAGGAGAAATGTTGAAGCAAGCGTCAGAGAGATGACT 1517
QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1518 TGGAAAGCCGAAATGAAATGGGTGCTTCAGGAGACATGAGGGCTGCCAACCTTTGGCCAA 1577
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAAGAACAGCCCTGGCCTTGTCCCTGACGGCCGACC 180
Db 1578 GCCCTCTTGTGATTAAAGCACACTAAGAAAGAAATAGCCCTGCCTTGTCTTGCACAGCTGACC 1637
QY 181 AGATGTCAGTGCCTTGTGATGCTGAGCCCCCCCCATACTCTATTCGAGTATGATCCTA 240
Db 1638 AGATGTCAGTGCCTTGTGATGCTGAGCCCCCCCCATACTCTATTCGAGTATGATCCTT 1697
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACTGGCAGACAGGGAGC 300
Db 1698 CTAGACCCCTTCAGTGAAGCCCTCAATGATGGGCTTATTGACCAACTAGCAGATAGGGAGC 1757
QY 301 TGGTTCACATGATCAACTGGGCGAGAGGGTGCCAG 336
Db 1758 TGGTTCATATGATCAACTGGGCGAAGAGAGTGCCAG 1793
RESULT 6
AK077236 2591 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
DEFINITION RIKEN full-length enriched library, clone:5031409B20
product:estrogen receptor 1 (alpha), full insert sequence.
ACCESSION AK077236
VERSION AK077236.1 GI:26097248
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913

QY 29 AATGTTGAAACACAGCG-CCAGAGAGATGATGGGAGGGCAGGGTGAAAGTGGGTCTG 87
|||||
Db 2 AATGTTTAAGCCCAAGCGTCAAAAAGATAA CTTGGAAGGCCAAAATGAAATGGGTGCTT 61
|||||
QY 88 CTGG-AGACATGAGAGCTGCCAACCTTTGGCCAAAGCCCGCTCATGATCAAAACGCTTAAG 146
|||||
Db 62 CAGGAAGACATGAGGGTTGCCAACCTTTGGCCAAAGCCCTCTTGTGATTAAGCACACTAAG 121
|||||
QY 147 AAGAACAGCCTGGCCCTTGTCCCTGACGGCCGACAGATGGTCAGTGCCTTGTGGATGCT 206
|||||
Db 122 AAGAATAGCCCTGCTTGTCCCTTGACAGCTGACCCAGATGGTCAGTGCCTTGTGGATGCT 181
|||||
QY 207 GAGCCCCCATACTCTATTCGAGTATGATCCTACAGACCCCTTCAGTGAAGCTTCGATG 266
|||||
Db 182 GAACCGCCCATGATCTATTCTGAATATGATCCTTCTAGACCCCTTCAGTGAAGCCTCAATG 241
|||||
QY 267 ATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTACATGATCAACTGGCGGAAG 326
|||||
Db 242 ATGGGCTTATTGACCAACCTAGCAGATAGGGAGCTGGTTTCATATGATCAACTGGGCAAAG 301
|||||
QY 327 AGGTGCCAG 336
|||||
Db 302 AGAGTGCCAG 311
|||||
RESULT 14
AU041270/c
LOCUS
DEFINITION
AU041270
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 576)
Ko,M.S.H., Kitchen,J.R., Wang,X., Threat,T.A., Sun,T.,
Depalma,G.E., Liang,Y., Kargul,G.J., Sharara,R., Lim,M.K. and
Doi,H.
Systematic analyses of genes expressed in 4-cell mouse embryo (The
ERATO/Doi Project at Wayne State University)
Unpublished (1998)
Contact: Hirofumi Doi
Doi Bioasymetry Project, ERATO
Japan Science and Technology Corporation (JST)
WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
Email: hdo@bio.jst.go.jp.
Location/Qualifiers
1. 576
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="J0828F12"
/dev_stage="four-cell-embryo"
/clone_lib="Mouse four-cell-embryo cdna"
ORIGIN
Query Match 50.8%; Score 170.8; DB 1; Length 576;
Best Local Similarity 81.6%; Pred. No. 2.6e-37;
Matches 209; Conservative 0; Mismatches 44; Indels 3; Gaps 1;
QY 81 GGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAAGCCCGCTCATGATCAAAACGC 140
|||||
Db 571 GGNNCCITCGGAGACATGAGGGCTGCCAACCTTTGGCCAAACCCCTCTTGTGATTA---GC 515
|||||
QY 141 TCTAAGAGAACAGCCTGGCCCTTGTCCTTGACGGCCGACAGATGGTCAGTGCCTTGTG 200
|||||
Db 514 ACAC TAGAGAATACCCCTCCCTTGTCCTTAACAGCTGACCAAGATGGTCAGTGCCTTGTNG 455
|||||
QY 201 GATGCTGAGCCCCCCATACTCTATTCCGAGTATGATCCTACCAGACCCCTTCAGTGAAGCT 260
|||||

Db 454 AATGCTGAACCGCCCATGATCTATTCTGAATATGATCCTTCTAGACCCCTTCAGTGAAGCC 395
|||||
QY 261 TCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTACATGATCAACTGG 320
|||||
Db 394 TCAATGATGGGCTTATTGACCAACCTAGCAGATAGGGAGCTGGTTCATATGATCAACTGG 335
|||||
QY 321 GCGAAGAGGGTGCCAG 336
|||||
Db 334 GCAAAGAGAGTGCCAG 319
|||||
RESULT 15
BB097271
LOCUS
DEFINITION
BB097271 RIKEN full-length enriched, 12 days embryo, embryonic body
between diaphragm region and neck Mus musculus cdna clone
9430058N13 3', mRNA sequence.
BB097271
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 316)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cdna. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsuunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cdna cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
Location/Qualifiers
1. 316
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9430058N13"
/tissue_type="embryonic body between diaphragm region and
neck"
/dev_stage="12 days embryo"

/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo,
embryonic body between diaphragm region and neck"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 370.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATAATTAATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI."

ORIGIN

Query Match	35.2%;	Score 118.4;	DB 2;	Length 316;
Best Local Similarity	83.8%;	Pred. No. 1.5e-22;		
Matches 134;	Conservative 0;	Mismatches 26;	Indels 0;	Gaps 0;
Qy	177	GACCAGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATCTATTCCGAGTATGAT	236	
Db	2	GACCAGATGATCAGTGCCTCATTTGGATCCTGGACCGCCCATTTATCTTTCTGGATATGTT	61	
Qy	237	CCTACCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGG	296	
Db	62	CCTTCTTGACCCCTTCAGTGAACCCCTCAATGATGGGCTTATTGACCAACCTAGCAGACAGG	121	
Qy	297	GAGCTGGTTCATGATCAACTGGGCGAAGAGGGGTGCCAG	336	
Db	122	GAGCTGGTTCATGATCAACTGGGCGAAGAGAGTGCAG	161	

Search completed: November 3, 2004, 09:43:36
Job time : 5803.75 secs

Search

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 04:55:34 ; Search time 2096.38 Seconds
(without alignments)
7579.407 Million cell updates/sec

Title: US-09-933-267A-1_COPY_306168_306503
Perfect score: 336
Sequence: 1 ggatacgaagaccgaaga.....ctggcggaagggtgccag 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	336	100.0	1223	6	AX066401	AX066401 Sequence
2	336	100.0	1237	9	HSESRI3	AF123496 Homo sapi
3	336	100.0	1374	6	AX066402	AX066402 Sequence
4	336	100.0	1770	6	AX824420	AX824420 Sequence
5	336	100.0	1788	6	AX411751	AX411751 Sequence
6	336	100.0	1803	6	AX824417	AX824417 Sequence
7	336	100.0	2092	6	I08538	I08538 Sequence 1
8	336	100.0	2092	6	AX474709	AX474709 Sequence
9	336	100.0	2092	9	HUMERMCF	M12674 Human estro
10	336	100.0	2106	9	HSU47678	U47678 Human 80 kD
11	336	100.0	2220	6	BD016996	BD016996 Fused est
12	336	100.0	2322	6	A42099	A42099 Sequence 1
13	336	100.0	2532	6	BD187708	BD187708 Specific
14	336	100.0	5439	9	HSM807087	BX640939 Homo sapi
15	336	100.0	6450	6	AR153585	AR153585 Sequence
16	336	100.0	6450	6	AR225684	AR225684 Sequence
17	336	100.0	6450	6	AR397400	AR397400 Sequence
18	336	100.0	6450	6	AX587952	AX587952 Sequence
19	336	100.0	6450	9	HSER	X03635 Homo sapien

20	336	100.0	6610	12	AF061181	AF061181 Mammalian
21	336	100.0	6833	6	AX128349	AX128349 Sequence.
22	336	100.0	6956	6	AX128348	AX128348 Sequence
23	336	100.0	7038	6	AX128342	AX128342 Sequence
24	336	100.0	11365	6	AX512493	AX512493 Sequence
25	336	100.0	165237	6	AX232505	AX232505 Sequence
26	336	100.0	299637	9	AY425004	AY425004 Homo sapi
27	336	100.0	349980	6	AX232503	AX232503 Sequence
28	336	100.0	349980	6	AX453703	AX453703 Sequence
29	334.4	99.5	139687	9	AL590993	AL590993 Human DNA
30	333	99.1	8566	6	AR397407	AR397407 Sequence
31	332.8	99.0	4963	6	AR029418	AR029418 Sequence
32	332.8	99.0	4963	6	I15368	I15368 Sequence 16
33	332.8	99.0	4963	6	I56760	I56760 Sequence 9
34	325	96.7	6639	6	AX128351	AX128351 Sequence
35	325	96.7	6695	6	AX128347	AX128347 Sequence
36	325	96.7	6695	6	AX128353	AX128353 Sequence
37	325	96.7	6695	6	AX128354	AX128354 Sequence
38	325	96.7	6801	6	AX128355	AX128355 Sequence
39	325	96.7	6818	6	AX128346	AX128346 Sequence
40	325	96.7	6900	6	AX128341	AX128341 Sequence
41	320.4	95.4	392	9	AF120105	AF120105 Homo sapi
42	288	85.7	1788	4	SSESTREC	Z37167 S.scrofa mR
43	280	83.3	1966	4	OAESTREC	Z49257 O.aries mRN
44	280	83.3	2327	4	AY538775	AY538775 Bos tauru
45	267.2	79.5	3746	4	AF124093	AF124093 Equus cab

ALIGNMENTS

RESULT 1
AX066401
LOCUS AX066401 1223 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 1 from Patent WO0100823.
ACCESSION AX066401
VERSION AX066401.1 GI:12544111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gannon,F., Denger,S. and Flouriot,G.
TITLE Novel isoforms of the human estrogen receptor-g(a)
JOURNAL Patent: WO 0100823-A 1 04-JAN-2001; -
EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
FEATURES
Location/Qualifiers
source 1..1223
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match	100.0%;	Score 336;	DB 6;	Length 1223;
	Best Local Similarity	100.0%;	Pred. No. 2.1e-85;		
	Matches 336;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGATACGAAAAGACCGAAGAGAGAGGAGAGATGTTGAAACACAAAGCCGAGAGATGATG	60		
Db	196	GGATACGAAAAGACCGAAGAGAGGAGAGATGTTGAAACACAAAGCCGAGAGATGATG	255		
QY	61	GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	120		
Db	256	GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA	315		
QY	121	GCCCGCTCATGATCAACCGTCTTAAGAAGAACAGCGCTGGCTTGTCCCTGACGGCCGACC	180		
Db	316	GCCCGCTCATGATCAACCGTCTTAAGAAGAACAGCGCTGGCTTGTCCCTGACGGCCGACC	375		
QY	181	AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA	240		
Db	376	AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA	435		

QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
|||||
Db 436 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 495
|||||
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
|||||
Db 496 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 531
|||||
RESULT 2
HSESRI3
LOCUS HSESRI3 1237 bp DNA linear PRI 07-APR-2000
DEFINITION Homo sapiens estrogen receptor alpha (ESR1) gene, exon 4.
ACCESSION AF123496
VERSION AF123496.1 GI:5821720
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Schubert,E.L., Lee,M.K., Newman,B. and King,M.C.
TITLE Single nucleotide polymorphisms (SNPs) in the estrogen receptor
gene and breast cancer susceptibility
JOURNAL J. Steroid Biochem. Mol. Biol. 71 (1-2), 21-27 (1999)
MEDLINE 20084372
PubMed 10619354
REFERENCE
AUTHORS Schubert,E.L., Lee,M.K. and King,M.-C.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1999) Medical Genetics, University of Washington,
1959 NE Pacific Street, Seattle, WA 98195-7720, USA
FEATURES
source
1. .1237
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q25"
350. .685
/gene="ESR1"
/number=4
exon
ORIGIN
Query Match 100.0%; Score 336; DB 9; Length 1237;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGATGTTGAAACACAAGCGCCAGAGATGATG 60
|||||
Db 350 GGATACGAAAAGACCGAAGAGGAGGGAGATGTTGAAACACAAGCGCCAGAGATGATG 409
|||||
QY 61 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
|||||
Db 410 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 469
|||||
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
|||||
Db 470 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 529
|||||
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
|||||
Db 530 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 589
|||||
QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
|||||
Db 590 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 649
|||||
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
|||||
Db 650 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 685
|||||

RESULT 3
AX066402
LOCUS AX066402 1374 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 2 from Patent WO0100823.
ACCESSION AX066402
VERSION AX066402.1 GI:12544112
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gannon,F., Denger,S. and Flourirot,G.
TITLE Novel isoforms of the human estrogen receptor_g(a)
JOURNAL Patent: WO 0100823-A 2 04-JAN-2001;
EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
FEATURES
source
1. .1374
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 336; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAAGACCGAAGAGGAGGGAGAGATGTTGAAACACAAGCGCCAGAGATGATG 60
|||||
Db 347 GGATACGAAAAGACCGAAGAGGAGGGAGAGATGTTGAAACACAAGCGCCAGAGATGATG 406
|||||
QY 61 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
|||||
Db 407 GGGAGGGCAGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 466
|||||
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
|||||
Db 467 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 526
|||||
QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
|||||
Db 527 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 586
|||||
QY 241 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
|||||
Db 587 CCAGACCCCTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 646
|||||
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
|||||
Db 647 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 682
|||||
RESULT 4
AX824420
LOCUS AX824420 1770 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 6 from Patent WO03070975.
ACCESSION AX824420
VERSION AX824420.1 GI:39750420
KEYWORDS
SOURCE
ORGANISM synthetic construct
synthetic construct
artificial sequences.
REFERENCE
AUTHORS Verdonk,G.N., Dijkema,R.N. and Schoonen,W.G.
TITLE An estrogen receptor alpha construct for use in a yeast 2-hybrid
JOURNAL assay
Patent: WO 03070975-A 6 28-AUG-2003;
Akzo Nobel N.V. (NL)
FEATURES
source
1. .1770
/organism="synthetic construct"

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion DNA sequence page 17-19"

ORIGIN

Query Match 100.0%; Score 336; DB 6; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 743 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGAGATGATG 802

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 803 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 862

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 863 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 922

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 923 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 982

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 983 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1042

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1043 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1078

RESULT 5
AX4111751
LOCUS AX4111751 1788 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0228175.
ACCESSION AX4111751
VERSION AX4111751.1 GI:21444272
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chambon,P. and Metzger,D.
TITLE Transgenic mouse for targeted recombination mediated by modified Cre-er
JOURNAL Patent: WO 0228175-A 1 11-APR-2002;
ASS POUR LE DEV DE LA RECH (FR)
FEATURES
source Location/Qualifiers
1..1788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1788
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35258.1"
/db_xref="GI:21444273"
/translation="MTMTLHTKASGMALLHQIQNELEPLNRPLKIPLERPLGEVYL
DSSKPAVNYPEGAAEFNAAAANAQVGTGLPYGPGEAAAFSGNGLGFPPLNS
VSPSPLMLLHPPQLSPFLQPHGQQVPIYLENPSGYTVREAGPPAFYRPNSDNRQ
GRERLASTNDKGSMAVESAKETRYCAVNDYASGYHYGVWSCEGCKAFFKRSIQHND
YMCPATNQCTIDKNRRKSCQACRLKCYEVGMKGGIRKDRGGRMKHKRQDDGEG
RGEVGSAGDMRAANLWPSPLMIKRSKNSLALSLTDQMVSAALLDAEPILYSEYDPT
RPFSEASMMGLLTNLADRELVMINWAKRVPGFVLDLTHDQVHLLCEAWLEILMIGLV
WRSMEHPGKLLFAPNLLDRNQKCVGEWVEIFDMLLATSSFRFMNVLQGEFVCLKS
IILLNSGYVTFLSSTLKSLEEKDHIHRLVDKITDITLHLMKAGLTLQQQHQRLAQLL
LILSHIRHMSNKGMEHLYSMCKNVVPLYDLLLLLEMLDAHRLHAPTSRGGASVEETDQS
HLATAGSTSSHSLSQKYIITGEAEGFPATV"

ORIGIN

Query Match 100.0%; Score 336; DB 6; Length 1788;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGATGATG 60
Db 761 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 821 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 881 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1060

QY 301 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGCGAAGAGGGTGCCAG 1096

RESULT 6
AX824417
LOCUS AX824417 1803 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 3 from Patent WO03070975.
ACCESSION AX824417
VERSION AX824417.1 GI:39750419
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Verdonk,G.N., Dijkema,R.N. and Schoonen,W.G.
TITLE An estrogen receptor alpha construct for use in a yeast 2-hybrid assay
JOURNAL Patent: WO 03070975-A 3 28-AUG-2003;
Akzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..1803
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion DNA p 14-16"

ORIGIN

Query Match 100.0%; Score 336; DB 6; Length 1803;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGATGATG 60
Db 776 GGATACGAAAGACCGAAGAGAGGGGAGAATGTTGAAACACAAGCGCCAGAGATGATG 835

QY 61 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 836 GGGAGGGCAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 895

QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 180
Db 896 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCCGACC 955

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 956 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTA 1015

QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1016 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1075
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1076 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1111

RESULT 7
I08538
LOCUS I08538 2092 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8705049.
ACCESSION I08538
VERSION I08538.1 GI:588753

KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2092)
AUTHORS Shine,J.
TITLE EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
JOURNAL Patent: WO 8705049-A 1 27-AUG-1987;

FEATURES
source Location/Qualifiers
1..2092
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 336; DB 6; Length 2092;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGAGATGATG 60
Db 1053 GGATACGAAAAGACCGAAGAGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGAGATGATG 1112

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1113 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172

QY 121 GCGCGTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGCGCGACC 180
Db 1173 GCGCGTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGCGCGACC 1232

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGGAGTATGATCCTA 240
Db 1233 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGGAGTATGATCCTA 1292

QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1293 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1352

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1353 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1388

RESULT 8
AX474709
LOCUS AX474709 2092 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 9 from Patent WO0222882.
ACCESSION AX474709
VERSION AX474709.1 GI:22214043

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Stanton,M., Epstein,D. and Hamaguchi,N.
TITLE Target activated nucleic acid biosensor and methods of using same

JOURNAL Patent: WO 0222882-A 9 21-MAR-2002;
FEATURES Archemix Corporation (US)
source Location/Qualifiers
1..2092
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 336; DB 6; Length 2092;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGAGATG 60
Db 1053 GGATACGAAAAGACCGAAGAGGAGGAGGAATGTTGAAACACAAAGCGCCAGAGAGATG 1112

QY 61 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1113 GGGAGGCGAGGGGTGAAGTGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1172

QY 121 GCGCGTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGCGCGACC 180
Db 1173 GCGCGTCATGATCAAAACGCTCTAAGAAGAACAGCGCTGGCCTTGTCCTGACGCGCGACC 1232

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGGAGTATGATCCTA 240
Db 1233 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCCATACTCTATTCGGAGTATGATCCTA 1292

QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1293 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1352

QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1353 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1388

RESULT 9
HUMERMCF
LOCUS HUMERMCF 2092 bp mRNA linear PRI 08-NOV-1994
DEFINITION Human estrogen receptor mRNA, complete cds.
ACCESSION M12674
VERSION M12674.1 GI:182192
KEYWORDS estrogen receptor.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2092)
AUTHORS Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and Shine,J.

TITLE Sequence and expression of human estrogen receptor complementary DNA

JOURNAL Science 231 (4742), 1150-1154 (1986)
MEDLINE 86122927
PUBMED 3753802

COMMENT Original source text: Human breast cancer cell line MCF-7, cDNA to mRNA, clone OR8.

FEATURES
source Location/Qualifiers
1..2092
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="6q24-q27"
/cell_line="MCF-7"
/cell_type="Mammary Carcinoma"
/lab_host="Prokaryotae"
/tissue_lib="lambda gt11"
1..2092
/gene="ESR"
<1..>2092
/gene="ESR"

gene
mRNA

ORIGIN

```
1 bp upstream of EcoRI site.
HLTAGTSSSHSLQKYIITGEAEGFPATV"
LLLSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLLLEMLDAHRLHAPT"SRGGASVEETDQS
IIILNSGVYITFLSSTILKSLEEKDTHIRVLDKITDITLHLMKAGATLQQQHQLAQLL
WRSMHEPVKLLFAPNLLLDNRQKCKVEGWEIIFDMLLATSSRFNMNLQGEFVCLKS
RPFSEASMGLLTNLADRELVHMTINWAKRVPGFVDLTLDHQVHLLDFCAWLEILIMIGLV
RGEVTSAGDMRAANLWPSPLMIKRKKNSLALSTADQMVSALLDAEPPILYSEYDPT
YMCPTATNQCTIDKNRRKSCQACRLKCYEVGMKGGIKORRRGRLMKHKRQRDDGEG
GRERLASTNDKGSMAVESAKETRYCAVGNDYASGYHYGVWSCGCKAFFKRSIQGHND
VSPFLMLLHPPQLSPFLQPHGQQVYYLENEPSGYTVREAGPAPYRPNSDNRRQGG
DSSKPAVINYPEGAAAYEFNAAAAANAQVYQTLPGYPGSEAAAFNGSLGGFFPPLNS
/translacion="MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGVEVYL
/db_xref="GDB:G00-119-120"
/db_protein_id="AAA52399.1"
/db_xref="GI:182193"
/codon_start=1
/note="estrogen receptor"
/gene="ESR"
293..2080
/product="ER mRNA"
```

Center, Northwestern University Medical Center, 303 E. Chicago Ave., Olson 8258, Chicago, IL 60611, USA

Location/Qualifiers

1. .2106

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="6"

/map="6q25"

/cell_line="MCF-7"

/tissue_type="breast"

CDS

1. .2106

/note="80 kDa estrogen receptor with an exon 6 and 7 repeat; It is able to bind estrogen"

/codon_start=1

/product="80 kDa estrogen receptor"

/protein_id="AAB00115.1"

/db_xref="GI:1197855"

/translation="MTMTLHTKASGMALLHQIQNELEPLNRPQLKIPLERPLGEVYL
DSSKPAVINYPEGAAAYEFNAAAAAQVYGTGLPYGPGSEAAAFSGNGLGFPPLNS
VSPSLMLHPPQLSPFLQPHGQQVPYYLENEPSGYTVREAGPAFYRPNSDNRQQ
GRERLASTNDKGSMAVESAKETRYCAVCNDYASGYHYGVWSCGCKAFFKRSIQGHND
YMCBATNQCTIDKNRRKSCQACRLKCYEVMGKGIKRRRGERMLKHKRQRDDGEG
RGEVSGAGDMRAANLWSPPLMIKRSKNSLALSLTADQMVSAALLDAEPPILYSEYDPT
RPFSEASMMGLLTNLADRELHMINWAKRVPGFVDLTLDQVHLLCAWLEILMIGLV
WRSMHPVKLLFAPNLLDRNQKCVGEMVEIFDMLLATSSRFMRMNLQGEFVCLKS
IILLNSGYTVFLSLLKSLKSEKDHHRVLDKITDTLIHLMAKAGLTQQQHQRLAQLL
LILSHIRHMNQKCVGEMVEIFDMLLATSSRFMRMNLQGEFVCLKSILLNSGYTV
FLSLTKSLEEKDHIHRVLDKITDTLIHLMAKAGLTQQQHQRLAQLLILSHIRHMS
NKGMHLYSMCKNVVPLYDLLLLLEMLDAHRLHAPTSRGGASVEETDQSHLATAGTSTS
HSLOXYITGEAEGFPAIV"

exon

1. .450

/number=1

exon

451. .642

/number=2

exon

643. .759

/number=3

exon

760. .1095

/number=4

exon

1096. .1233

/number=5

exon

1234. .1368

/number=6

exon

1369. .1551

/number=7

exon

1551. .1685

/note="exon 6"

exon

1686. .1868

/note="exon 7"

exon

1869. .2106

/number=8

ORIGIN

Query Match 100.0%; Score 336; DB 9; Length 2106;

Best Local Similarity 100.0%; Pred. No. 2.1e-85;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGAGGAGGAGAATGTTGAAAACACAAGCGGCAGAGAGATGATG 60

Db 761 GGATACGAAAAGACCGAAGAGGAGGAGGAGAATGTTGAAAACACAAGCGGCAGAGAGATGATG 820

QY 61 GGGAGGGCAGGGGTGAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120

Db 821 GGGAGGGCAGGGGTGAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 880

QY 121 GCCCGCTCATGATCAAACGCTCTAAGAAGAACAGCCTGGCCCTTGTCCCTGACGGCCGACC 180

Db 881 GCCCGCTCATGATCAAACGCTCTAAGAAGAACAGCCTGGCCCTTGTCCCTGACGGCCGACC 940

QY 181 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 240

Db 941 AGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCCTACTCTATTCCGAGTATGATCCTA 1000

QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACTGGCAGACAGGGAGC 300
Db 1001 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACTGGCAGACAGGGAGC 1060
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1061 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1096

RESULT 11
BD016996
LOCUS BD016996 2220 bp DNA linear PAT 27-AUG-2002
DEFINITION Fused estrogen receptor protein.
ACCESSION BD016996
VERSION BD016996.1 GI:22558172
KEYWORDS JP 2001258572-A/8.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 2220)
AUTHORS Matsui,K., Ishibashi,T. and Kawamura,Y.
TITLE Fused estrogen receptor protein
JOURNAL Patent: JP 2001258572-A 8 25-SEP-2001;
TOYOBO CO LTD

COMMENT PN JP 2001258572-A/8
PD 25-SEP-2001
PF 23-MAR-2000 JP 2000082034
PI KAZUHIRO MATSUI,TAKUYA ISHIBASHI,YOSHIHISA KAWAMURA PC
C12N15/09,C07K14/00,C07K14/72,C07K19/00,C12N1/21,C12P21/02, PC
G01N33/15,
PC
G01N33/50,G01N33/53,G01N33/566// (C12N1/21,C12R1:19), (C12P21/02, PC
C12R1:19),
PC C12N15/00
CC Strandedness: Both;
CC Topology: Both; Location/Qualifiers.
FH Key Location/Qualifiers.

FEATURES
source
1..2220
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 336; DB 6; Length 2220;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 1196 GGATACGAAAAGACCGAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGATGATG 1255
QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1256 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1315
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 180
Db 1316 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 1375
QY 181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1376 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 1435
QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1436 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1495
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1496 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1531

RESULT 12

A42099
LOCUS A42099 2322 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 1 from Patent WO9500555.
ACCESSION A42099
VERSION A42099.1 GI:2297591
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE 1 (bases 1 to 2322)
AUTHORS Stewart,F.
TITLE REGULATION OF SITE-SPECIFIC RECOMBINATION BY SITE-SPECIFIC
RECOMBINASE/NUCLEAR RECEPTOR FUSION PROTEINS Regulation of
site-specific recombination by site-specific recombinase/nuclear
receptor fusion proteins

JOURNAL Patent: WO 9500555-A 1 05-JAN-1995;
EUROP MOLECULAR BIOLOGY LAB BM (DE)
COMMENT Other publication AU 7227994 950117.
FEATURES Location/Qualifiers

source
1..2322
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..2322
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA02659.1"
/db_xref="GI:2297592"
/translation="MPQFDILCKTPPKVLVRQVFERFERPSGEKIALCAAELTYLCWM
ITHNGTAIKRATFMSYNTIISNSLSDIVNKSLOQKYKTQKATILEASLKKLIPAWEF
TIIPYQKHQSDITDIVSSLQLFESSEADKGNSHSKMLKALLSEGESIWEITEK
ILNSEYTSRFTKTKLYQFLATFINGRFSDIKNVDPKSPKLVQNKYLGVIIOCL
VTETKTSVSRHIYFFSARGRIDPLVYLDEFNRNSEPLKRVNRTGNSSSNKQEYQLLK
DNLVRSYNKALKKNAPYSIFAIKNGPKSHIGRHLMTSFLSMKGLTELTVVGNWSDKR
SAEGSIRYPANNGIISOEVLVYSSYINRRI SVRSGMKGGIRKDRRGRMLKHKRQD
DGEGRVGSAGDMAANLWPSPLMIKRSKNLSALSALTADQMVSAILDAPPPILYSE
YDTPRPFSEASMMGLLTNLADRELVMINWAKRVPGFVDLTLDHQVHLLCEAWLEILM
IGLVWRSMHPVKLLFAPNLLDRNQKCEGVMEIFDMLLATSSRFRMMNLQGEFV
CLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLOQHQRL
AQLLLILSHIRHMSNKGMEHLYSMCKKNVVPPLYDLLLEMLDAHLHAPT SRGGASVEE
TDQSHLATAGTSSSHLSQLQYYITGEAEGFPATV"

CDS

ORIGIN

Query Match 100.0%; Score 336; DB 6; Length 2322;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATACGAAAAGACCGAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 60
Db 1295 GGATACGAAAAGACCGAAGAGGGGAGAAATGTTGAAACACAAGCGCCAGAGAGATGATG 1354
QY 61 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db 1355 GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1414
QY 121 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 180
Db 1415 GCCCGCTCATGATCAAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCTGACGGCGGACC 1474
QY 181 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 240
Db 1475 AGATGGTCAGTGCCTTGTTGGATGCTGAGCCCCCCCATACTCTATTCCGAGTATGATCCTA 1534
QY 241 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db 1535 CCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1594
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db 1595 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1630

RESULT 13
BD187708
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BD187708
Specific cleavage method of fusion protein.
BD187708
BD187708.1 GI:32997447
JP 2003024095-A/1.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2532)
Nishiya,Y. and Oka,M.
Specific cleavage method of fusion protein
Patent: JP 2003024095-A 1 28-JAN-2003;
TOYOBO CO LTD
OS Homo sapiens (human)
PN JP 2003024095-A/1
PD 28-JAN-2003
PF 10-JUL-2001 JP 2001209476
PI YOSHIAKI NISHIYA,MASANORI OKA
PC Cl2P21/06//Cl2N15/09,Cl2N15/00
CC Specific cleavage method of fusion protein
FH Key
FT source
FT Location/Qualifiers
1. .2532
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
Location/Qualifiers
1. .2532
Query Match 100.0%; Score 336; DB 6; Length 2532;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAGAGGGGAGAAATGTTGAACACAAAGCGCCAGAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1505 GGATACGAAAGACCGAGAGAGGGGAGAAATGTTGAACACAAAGCGCCAGAGAGATGATG 1564
QY 61 GGGAGGGCAGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1565 GGGAGGGCAGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 1624
QY 121 GCCCGCTCATGATCAACCGCTCTAAGAAGAAGACAGCCTGGCCTTGTCCCTGACGCCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1625 GCCCGCTCATGATCAACCGCTCTAAGAAGAAGACAGCCTGGCCTTGTCCCTGACGCCGACC 1684
QY 181 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCATCTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1685 AGATGGTCAGTGCCCTGTTGGATGCTGAGCCCCCCATCTCTATTCCGAGTATGATCCTA 1744
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1745 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 1804
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1805 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 1840
RESULT 14
HSM807087
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
HSM807087
Homo sapiens mRNA; cDNA DKFZp686N23123 (from clone DKFZp686N23123);
complete cds.
BX640939
BX640939.1 GI:34365270
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5439)
Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
The German Human cDNA Consortium
Direct Submission
Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686N23123) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1. .5439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="6q25.1"
/clone="DKFZp686N23123"
/tissue type="human uterus"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B;_sites SfiIA + SfiIB"
/dev_stage="adult"
1. .5439
/gene="DKFZp686N23123"
234. .1166
/gene="DKFZp686N23123"
/note="estrogen receptor alpha, differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45969.1"
/db_xref="GI:34365271"
/translation="MAMESAKETRYCAVCNDYASGYHYGVWSCGCKAFFKRSIQGHN
DYMCPTNQCTIDKNRKSQCACRLKCYEVGMKGIKDRRGGRMLKHKRQRDDGE
GRGEVGSAGDMRAANLWPSPLMIKRSKNSLALSLTADQMVSAALLDAEPPILYSEYDP
TRPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLDQVHLLLECAWLEILMIGL
VWRSMHPKLLFAPNLLDRNQKCVGEMVEIFDMLLATSSRFMMNLQGEFVCLK
SILLNSGISHVAKRILNLHPKIFGNKWFPRV"
5416
polyA_site
/gene="DKFZp686N23123"
ORIGIN
Query Match 100.0%; Score 336; DB 9; Length 5439;
Best Local Similarity 100.0%; Pred. No. 2.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATACGAAAGACCGAAGAGGAGGGAGATGTTGAACACAAAGCGCCAGAGATGATG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
475 GGATACGAAAGACCGAAGAGGAGGGAGATGTTGAACACAAAGCGCCAGAGATGATG 534
QY 61 GGGAGGGCAGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
535 GGGAGGGCAGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAA 594
QY 121 GCCCGCTCATGATCAACCGCTCTAAGAAGAAGACAGCCTGGCCTTGTCCCTGACGCCGACC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
595 GCCCGCTCATGATCAACCGCTCTAAGAAGAAGACAGCCTGGCCTTGTCCCTGACGCCGACC 654
QY 181 AGATGGTCAGTGCCCTTGTGGATGCTGAGCCCCCCATCTCTATTCCGAGTATGATCCTA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
655 AGATGGTCAGTGCCCTTGTGGATGCTGAGCCCCCCATCTCTATTCCGAGTATGATCCTA 714
QY 241 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
715 CCAGACCCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGC 774
QY 301 TGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAG 336

Search completed: November 3, 2004, 07:43:08
Job time : 2097.38 secs